2002

Using simulated annealing in geostatistics

Wesley Wells
Edith Cowan University

Edith Cowan University

Copyright Warning

Edith Cowan University

Faculty of Communications, Health and Science

School of Engineering and Mathematics

Using Simulated Annealing in Geostatistics.

Wesley Wells

Bachelor of Science (Mathematics) Honours.

August 2002

USE OF THESIS

The Use of Thesis statement is not included in this version of the thesis.

ABSTRACT

Simulation methods are now used extensively for estimation and prediction in mining and petroleum industries and also in environmental management. In this thesis we describe the method of simulated annealing and examine in detail the GSLIB implementation algorithm SASIM.

In the context of two case studies involving both sample and exhaustive data sets, we demonstrate this algorithm and then investigate the effect on the outcome of varying the different algorithm parameters. We also consider the effect of varying the weighting given in the simulated annealing objective function to the reproduction of each of the sample histogram and semivariogram.

For the two case studies, we implement also the more commonly used method of Sequential Gaussian Simulation and compare the results with those obtained from Simulated Annealing.

DECLARATION

I certify that this thesis does not, to the best of my knowledge and belief:

(i) *incorporate without acknowledgement any material previously submitted for a degree or diploma in any institution of higher education;*

(ii) *contain any material previously published or written by another person except where due reference is made in the text; or*

(iii) *contain any defamatory material.*

29/08/2002

Wesley Wells

ACKNOWLEDGEMENTS

Firstly I would like to thank my supervisors Associate Professor Lyn Bloom and Dr Ute Mueller for their guidance, their support and their teachings, and for believing in me throughout the thesis process. Next I would like to thank the other mathematics lecturers Dr David McDougall, Geoff Comber and Dr Pender Pedler for their teachings and support throughout the Bachelor course.

In a personal sense, I would like to thank my family for their support while I complete university studies. I would also like to thank my classmates for their support and assistance throughout my course and honours year.

CONTENTS.

1. INTRODUCTION.

Geostatistics is an area of mathematics and statistics that allows us to study the spatial fluctuation of natural variables from one region in space to another. Examples of such variables include permeability in a porous medium, soil properties in a region and ore grades in a mineral deposit. Geostatistics can also be used to measure the concentrations of pollutants in a contaminated site, or for estimating the number of fish in a region (Rivoirard, et al., 2000). Geostatistics has been used to model real-world phenomena in fields such as mining, environmental management, petroleum exploration, and oceanography. There are several books that describe the techniques of geostatistics. Core texts used for this thesis include Goovaerts (1997), Deutsch and Journel (1998), Chiles and Delfiner (1999), and Isaaks and Srivastava (1989). The reader's attention is called to these texts and the many references to journal articles that they contain. For this thesis we adopt the notation used by Goovaerts (1997).

This thesis concentrates on simulation methods as used in geostatistics. Simulation allows us to estimate at unsampled locations by using the known values with algorithms to generate a number of realizations. There are many different simulation algorithms available (Deutsch and Journel, 1998; Armstrong and Dowd, 1994) but in this thesis we will concentrate on two. These are Sequential Gaussian Simulation and Simulated Annealing. There have been several articles written previously to compare different simulation methods including those by Gotway and Rutherford (1994) and Hegstad et al. (1994). We will compare the results from Simulated Annealing with those from Sequential Gaussian Simulation. Both Sequential Gaussian and Simulated Annealing are conditional in the sense that they honour the data at the sample locations.

Sequential Gaussian Simulation is examined first in Sections 4 and 5. In Section 4 we also examine simple kriging in preparation for its use in Sequential Gaussian Simulation. Ultimately the greater concentration of this thesis is on the background and the uses of Simulated Annealing, especially in geostatistics, in Section 6 and 7. There have been many journal articles written that examine

Simulated Annealing in geostatistics. These include Fang and Wang (1997), Carle (1997), Deutsch and Wen (1998), Goovaerts (1998), Li, et al. (2000) and Caers (2001). Other articles written on Simulated Annealing as a solver for optimisation problems include Aarts (1989), Davis (1990), Azencott (1991), Kalivas (1995), Rosenhouse-Dantsker and Osman (2000), Nolte and Schrader (2001) and Wang, et al. (2001).

To examine the simulation methods in geostatistics we will make use of the computer algorithms in GSLIB (Deutsch and Journel, 1998). For Simulated Annealing, there are several aspects that will be experimented with in the parameter file. They are (1) using only the semivariogram component, (2) using semivariogram and histogram components, (3) using different weights for the semivariogram component, (4) using a different number of lags, and (5) using different annealing schedules.

In Section 8, the selected Simulated Annealing realizations from these experiments will be compared with the realizations from Sequential Gaussian Simulation. The selected realizations from Simulated Annealing will be evaluated by observing how closely the histogram and semivariogram of the sample have been reproduced. The comparisons between these Simulated Annealing realizations and the Sequential Gaussian realizations will be made by using mean absolute deviation and mean square error.

To examine the performance of Simulated Annealing in geostatistics we need to use a sample data set. In this thesis we will apply Simulated Annealing to two different sample data sets, which are described in detail in Section 2. The first sample data set, *True150* consists of 150 observations on a 50 by 50 grid. This data set is positively skewed and isotropic. The second data set, *Berea100* contains 100 observations on a 40 by 40 grid. This data set is approximately normally distributed and anisotropic.

2. THE DATA SETS.

In order to explore in detail the Simulated Annealing algorithm SASIM, and to compare it with the Sequential Gaussian algorithm SGSIM, two sample data sets will be used. These are *True150* and *Berea100*, which contain observations taken from the exhaustive data sets *True* and *Berea* respectively. A description of each data set is given below.

2.1 *True150.*

True is a data set that comes with the GSLIB software package (Deutsch and Journel, 1998). It has the characteristics of a gold mineralisation data set. It consists of 2500 observations on a two-dimensional 50 by 50 grid, with a grid spacing of one mile in each direction. *True* contains two attribute variables, Primary and Secondary, but only data associated with Primary will be used in this thesis. *True150* is a random sample consisting of 150 observations taken from *True* and created for this thesis. *True150* can be viewed in Appendix A.1.

In Figure 2.1, we give the descriptive statistics for both the *True* and *True150* data sets.

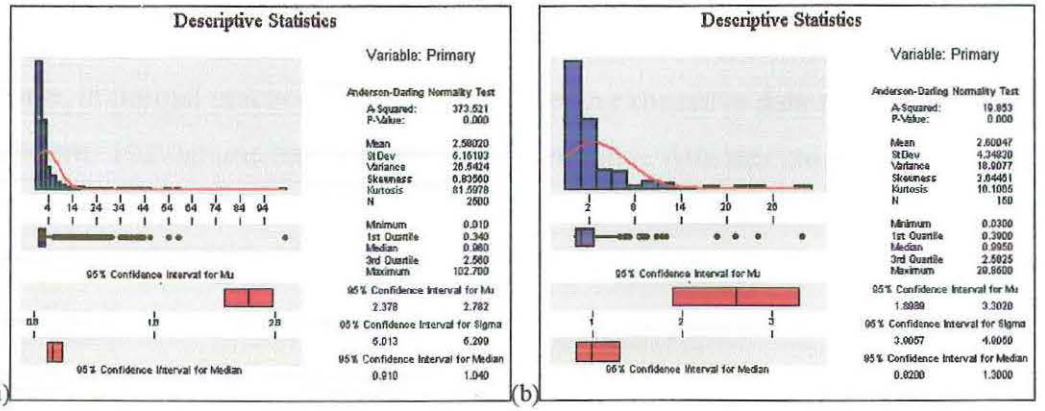

Figure 2.1: Graphical descriptive statistics for True (a), and True150 (b).

From the boxplot in Figure 2.1b, we can see that all but four observations in *True150* are less than the value 14. In Figure 2.1a we can see that the majority of values of *True* lie below 14. In fact 75% of the values are below 2.560 for *True*

and 2.582 for *True150*. From Figure 2.1b, we can see that the sample data are not normally distributed. This is further shown in the Anderson-Darling normality test where the p-value is zero for both exhaustive and sample data sets. Comparing the statistics for both, we first see that the differences between the means and medians of the exhaustive and sample data sets are relatively small. These differences are 0.02 and 0.023, for the mean and medians respectively. This is further supported by the 95% confidence intervals for the mean and median, since the mean and median of the exhaustive data set fall within the upper and lower limits derived from the sample set in Figure 2.1b. Because the mean and median are close and the two distributions are similar, we can conclude that the sample set is representative of the exhaustive set *True*.

There is however substantial difference in the ranges of *True* and *True150*. The range for the exhaustive set is 102.090, and the range for the sample set is 29.820. This difference is due to the extreme values seen in Figure 2.1a that are not found in the sample (Figure 2.1b). There is also a difference between the standard deviations. We can see in Figure 2.1b that the standard deviation of 5.152 for the exhaustive set falls outside the 95% confidence interval for the standard deviation of the sample data set.

In this thesis when we conduct simulation with our data sets we will compare the histogram of the realizations with the histogram of the sample. Of course, in normal practice we would not have an exhaustive data set with which to compare. Histograms for the sample and exhaustive data sets are shown in Figure 2.2. Each histogram is generated in Excel, where the size of each class is two. That is, the first column contains all observations with a value from 0 to 2, the next 2 to 4, and so on up to 104. But because the frequency of values is only one, they do not show on the histograms.

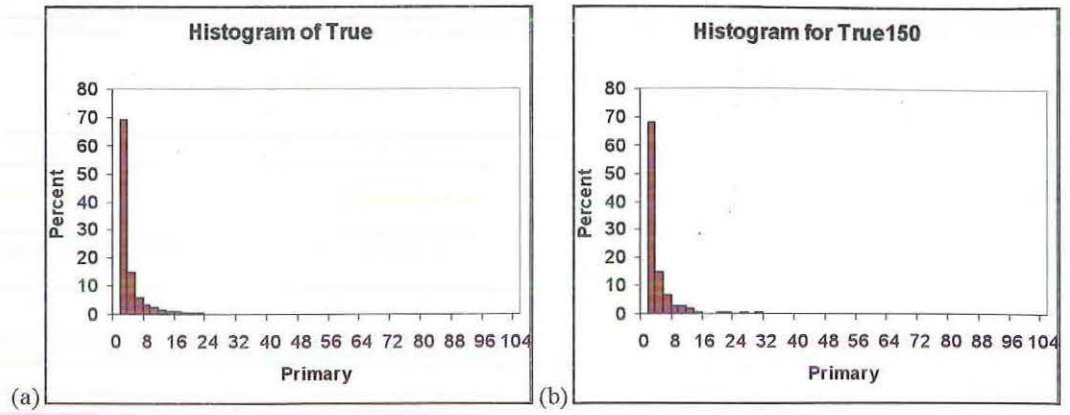

Figure 2.2: Histograms for the (a) exhaustive and (b) sample data.

The histograms in Figure 2.2 use a scale on the x-axis from 0 to 104. The value 104 is selected because the maximum value for the attribute Primary in *True* is 102.7 and for *True150* the maximum value is 29.85. Later in the thesis, the maximum value on the x-axis that we will use will be 50. Both histograms begin with a large number of values, approximately 70%, between zero and two, and then taper off to their maximum.

Figure 2.3 gives a mosaic map (a) and a post plot (b) of *True* and *True150* respectively. The levels indicated are those used by Deutsch and Journel (1998).

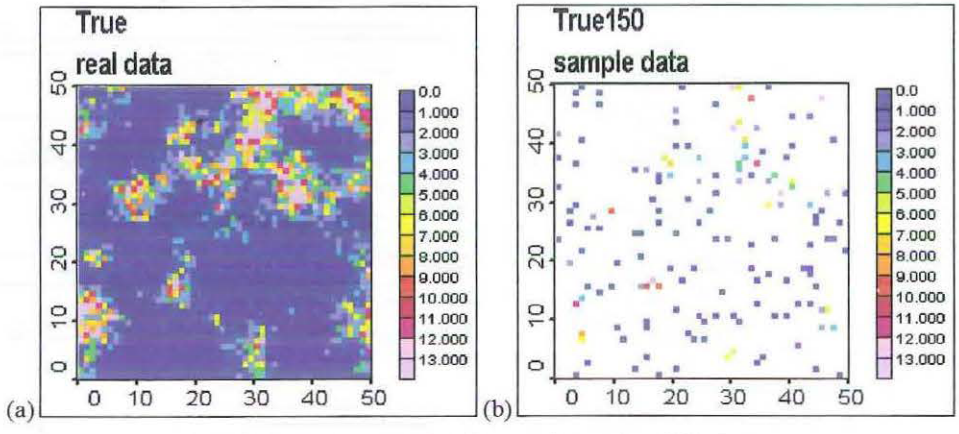

Figure 2.3: Mosaic map of (a) True and post plot of (b) True150.

The mosaic map for *True* shows the spatial distribution of the Primary values. We see that there are several regions with higher values greater than two surrounded by the low values (below two). Most of these higher values are found in the top right quadrant of the grid with other regions distributed in the other three quadrants. The lower values of Primary make up approximately 50% of the

locations. The post plot for the sample data set shows the sample locations as well as the value at each location. From this post plot, we can see that the majority of sample values are less than two.

2.2 *Berea100.*

Berea is a data set used in Chu, et al. (1991). It is a two-dimensional data set that is derived from measurements (in millidarcies) of air permeability taken from a 2 by 2 foot slab of Berea sandstone and is an example of an oil exploration data set. It consists of 1600 observations taken on a regular 40 by 40 grid, where the grid has 40 observations taken in north and east directions. The data set contains a fourth column for elevation, but all measurements are taken at a single elevation, so this column will be ignored. *Berea100* is a random sample data set consisting of 100 observations taken from *Berea* and created for this thesis. *Berea100* can be viewed in Appendix A.2.

In Figure 2.4, we give the graphical descriptive statistics for *Berea* in (a) and *Berea100* in (b).

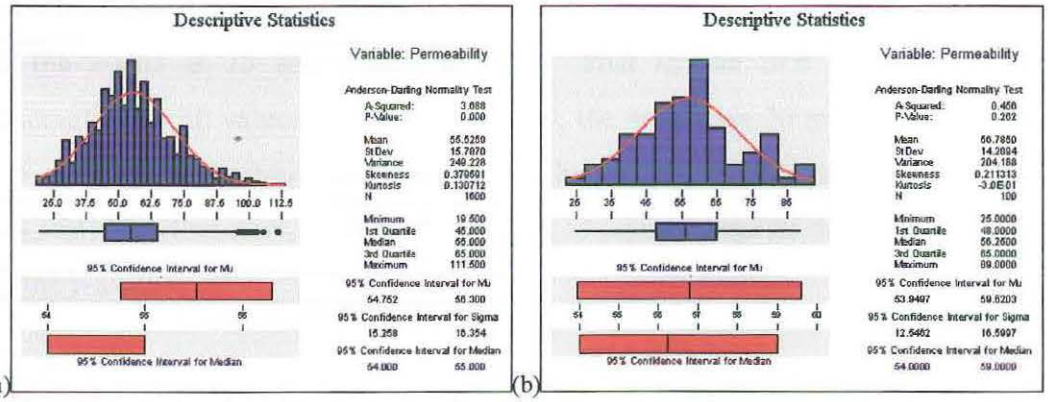

Figure 2.4: Graphical descriptive statistics for exhaustive data set (a), and sample data set (b).

From viewing the descriptive statistics in Figure 2.4b, it appears that the sample data are approximately normally distributed. This is confirmed by the Anderson-Darling normality test (p-value = 0.262) included for the sample data set. For the exhaustive data set, there is evidence that the data may not be normally distributed based on the Anderson-Darling normality test (p = 0.000).

Comparing the statistics for *Berea* and *Berea100* we can see that the mean and median for each are similar. The mean for the exhaustive and sample data are 55.526 and 56.785, respectively, giving only a relatively small difference of 1.259. The medians of the exhaustive and sample data are 55.000 and 56.250, respectively, so the difference is 1.25. The differences in the means and medians when we compare to the actual means and medians is relatively small and therefore they are not substantially different. This similarity between the respective means and medians can also be seen when we view the 95% confidence intervals for each. The mean and median of the exhaustive data set fall within the upper and lower confidence limits for mean and median of the sample data set derived in Figure 2.4b. Furthermore, the 95% confidence interval for the standard deviation in Figure 2.4b shows that the standard deviations for the exhaustive and sample data sets are similar. The ranges are different. The range for the sample data set is 64 whilst the range for the exhaustive data set is 92. This difference in the range could be explained by the presence of outliers in the exhaustive data that do not appear in the sample. We can conclude overall that the sample set is representative of *Berea*.

The histograms for *Berea* (a) and *Berea100* (b) are shown in Figure 2.5. Both histograms were generated in Excel and use classes of size five, beginning on the x-axis at 15 and ending at 115. That is, the first class counts all observations with values between 15 and 20, the next class 20 to 25 and so on to 115. Using the minimum and maximum values from the exhaustive data set sets this scale. Furthermore, each class is shown as a percentage of the whole data set on the y-axis.

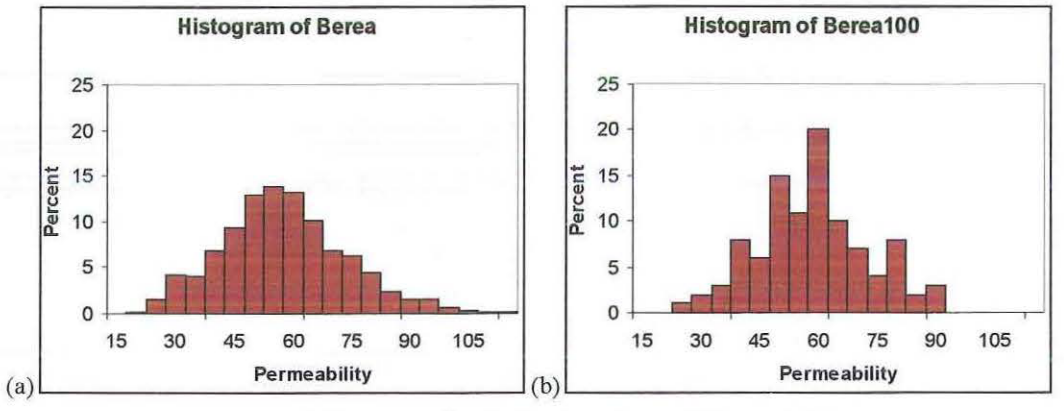

Figure 2.5: Histograms for the (a) exhaustive and (b) sample data.

The histograms in Figure 2.5 clearly show that the ranges of the exhaustive and sample data are different, which we saw earlier. We also see that in the histogram for *Berea100,* although it is a sample set taken from *Berea*, there are classes that have a larger percentage of values than seen in the exhaustive data set. This is expected. In Sections 5 and 7 we shall use the histogram of the sample for comparing with the histograms of the simulation maps.

In Figure 2.6, there is a mosaic map (a) of the exhaustive data set showing all the values and a post plot (b) of the sample data set containing 100 observations.

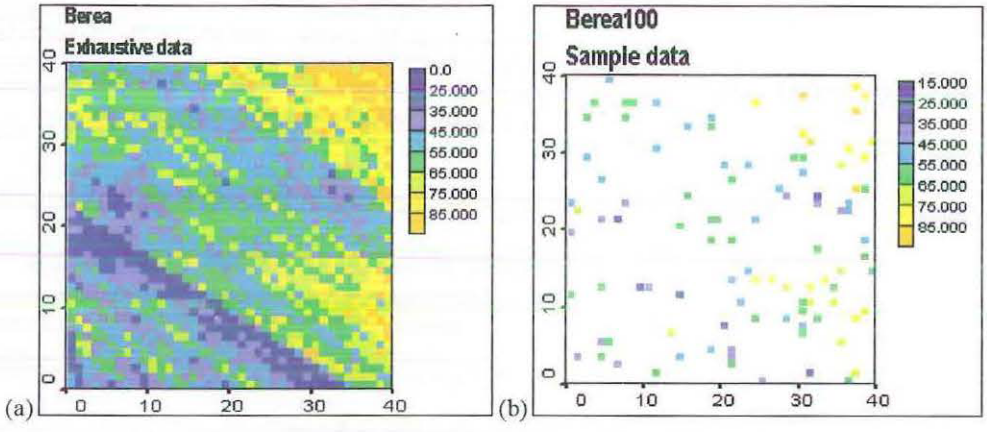

Figure 2.6: Mosaic map of Berea exhaustive set (a) and the sample set (b).

From viewing the mosaic map in (a), we can see that there is banding in the northwest-southeast direction and this is an early indication that there will be anisotropy. In the perpendicular direction, the values tend to increase with the higher values concentrated in the northeast corner of the study region. This increase is referred to as the trend. The post plot in (b) for the sample data set has captured several high values especially in the northeastern corner seen in the mosaic map in (a). Also, along the banding in (a), the sample appears to consist of the values between the first and third quartiles in Figure 2.4. These values could assist in keeping the banding in simulation.

2.3. Data transformation.

After the initial exploratory data analysis has been conducted, the next step is to carry out the variography. But before variography can proceed, a decision, based on the particular simulation algorithm to be used, as to whether or not to transform the data needs to be made. Sequential Gaussian simulation needs standard normal data and simulated annealing needs data that are not too skewed.

The method of transforming the data with positive skewness to normality for the Sequential Gaussian simulation algorithm SGSIM is normal score transformation. Normal score transformation (ϕ) begins by ranking the original data in ascending order, then computing the cumulative frequency of the datum, and finally the normal score of the datum is matched to the quantile of the standard normal cumulative distribution function (Goovaerts, 1997). That is, the original sample data are transformed into new data with a standard normal histogram. The new data are then referred to as normal scores. Since it is rare to find a data set in geostatistics that is standard normal, transformation is usually required. Both of our data sets, *True150* and *Berea100*, need to be transformed. In GSLIB, defined in Appendix B, a program and parameter file are available to us called NSCORE to transform the data appropriately.

For the Simulated Annealing algorithm SASIM, the method of deskewing is to use logarithm to base ten. To transform using this method we will use Minitab (Appendix B). We noted in Section 2.2 that *True150* is positively skewed whilst *Berea100* is approximately normally distributed. Therefore for Simulated Annealing, only *True150* needs to be log transformed.

3. VARIOGRAPHY.

3.1 Theory.

The objective of variography is to model the spatial continuity of an attribute under consideration. A random function $\{Z(\mathbf{u}), \forall \mathbf{u} \in A\}$ regards all known $z(\mathbf{u}_\alpha)$ and unknown values $z(\mathbf{u})$ as values of (possibly dependent) random variables $Z(\mathbf{u}_\alpha)$ and $Z(\mathbf{u})$, one random variable for each location $\mathbf{u}$ over the whole study region A. A random variable is distinguished as being either discrete or continuous. In this thesis we shall consider only continuous random variables. Examples of continuous random variables are metal concentrations and permeability. When expressing a cumulative distribution function of a continuous random variable, the following is used:

$$F(\mathbf{u}; z) = \text{Prob}\{Z(\mathbf{u}) \leq z\} \quad \forall z \tag{3.1}$$

Inference of this cumulative distribution function in (3.1) generally needs repetitive sampling of $z(\mathbf{u})$ to evaluate from experimental proportions. But in most cases only one sample is available at any particular location $\mathbf{u}$ which $z(\mathbf{u})$ is known. The idea behind the inference process is to exchange the unavailable replication at one location with one that is available elsewhere in space or time (Deutsch and Journel, 1998). This is the assumption of stationarity. A random function is therefore considered to be stationary of order two when the expected value $E\{Z(\mathbf{u})\}$ exists and is invariant within A, and the covariance $C(\mathbf{h})$ exists and depends only on the separation vector $\mathbf{h}$ (Goovaerts, 1997). The covariance measures the nature of the association between two random variables. The covariance is defined by:

$$C(\mathbf{h}) = E\{Z(\mathbf{u}) \cdot Z(\mathbf{u}+\mathbf{h})\} - E\{Z(\mathbf{u})\} \cdot E\{Z(\mathbf{u}+\mathbf{h})\} \tag{3.2}$$

In variography we use the variogram or semivariogram to find a model that is used to assist in finding reasonable simulation realizations. The variogram

is used to measure the spatial continuity of the data. The variogram $2\gamma(\mathbf{h})$ is defined by:

$$2\gamma(\mathbf{h}) = Var[Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h})] = E\{[Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h})]^2\} \qquad (3.3)$$

where the separation vector $\mathbf{h}$ is used to account for both distance $|\mathbf{h}|$ and direction. From (3.2) it can be seen that the variogram is measured as the expected squared difference between two random variables separated by vector $\mathbf{h}$. The semivariogram is one-half of the variogram and is generally modelled in preference to the variogram because we assume that $\gamma(\mathbf{h}) = \gamma(-\mathbf{h})$.

The semivariogram and covariance are related by $\gamma(\mathbf{h}) = C(0) - C(\mathbf{h})$. When $|\mathbf{h}|$ increases then $C(\mathbf{h})$ tends to zero and the semivariogram tends to $C(0)$.

When we begin searching for appropriate semivariograms from the sample data, we need to experiment with different lag spacings or distances. The aim is to find the smallest lag spacing that gives us a semivariogram that is smooth enough to model. The semivariograms we choose in this way to become our experimental semivariograms for further analysis.

The spatial continuity is said to be anisotropic if it depends on the direction. If it is independent of the direction then the semivariogram is isotropic. In order to determine if the data are isotropic a semivariogram surface is used. It is a map of experimental semivariogram values that assists in identifying the directions of maximum and minimum spatial continuity. Semivariogram surfaces can also help to find reasonable values for the number of lags and lag spacing that are used in semivariograms.

Once we have an experimental semivariogram, we need to find a suitable semivariogram model. A semivariogram model enables us to compute a semivariogram value at any separation vector. Trying to fit all lag points as well as possible so as to smooth out sample fluctuations generates a suitable semivariogram model. This model is then used when generating our simulation

realizations. There are several standard base models that are used and we will use a combination of two in this thesis. They are expressed in their isotropic form as:

- The nugget effect model defined by:

$$g(\mathbf{h}) = \begin{cases} 0 & \text{if } |\mathbf{h}| = 0 \\ 1 & \text{otherwise} \end{cases} \tag{3.4}$$

- The spherical model with a range a, defined by:

$$g(\mathbf{h}) = Sph\left(\frac{|\mathbf{h}|}{a}\right) = \begin{cases} 1.5 \cdot \frac{|\mathbf{h}|}{a} - 0.5 \cdot \left(\frac{|\mathbf{h}|}{a}\right)^3 & \text{if } |\mathbf{h}| \le a \\ 1 & \text{otherwise} \end{cases} \tag{3.5}$$

These models are bounded. For the nugget effect model, the sill is reached as soon as $|\mathbf{h}|$ is greater than zero and for the spherical model, the sill is reached at the distance a, called the actual range.

There are two types of anisotropy. They are:

- Geometric anisotropy where the sills are the same, but the ranges are different, and
- Zonal anisotropy where the sills are different.

The reader should also note that this is only a summary on the theory and further explanation can be obtained from Goovaerts (1997) and Deutsch and Journel (1998).

3.2 Variography for NSCORE transformed *True150*.

The semivariogram surfaces for the normal scores generated in GSLIB's program NSCORE, of *True* and *True150* are shown in Figure 3.1. The levels for each are identical, but the lag spacings are different. The lag spacing for the exhaustive data is set to one, whilst the lag spacing for the sample data is 2.5. These lag spacings are selected because the exhaustive set consists of values at all locations, hence a lag spacing equal to one, and the sample data has an average distance between the data pairs of about 2.5.

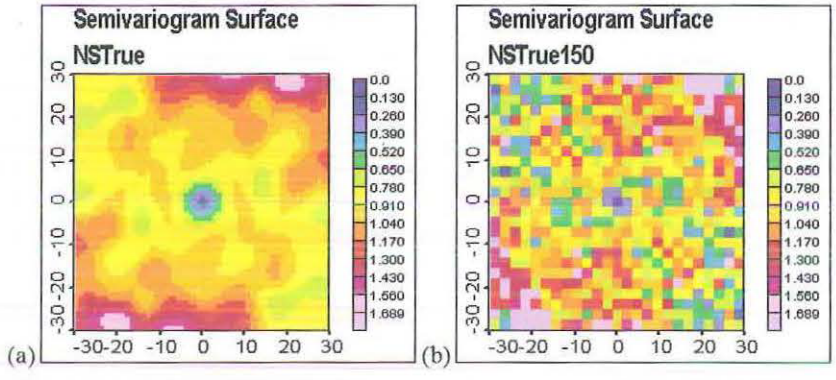

Figure 3.1: Semivariogram surfaces of normal scores for True (a) and True150 (b).

The semivariogram surface in Figure 3.1b for the sample data shows that the data are isotropic. Using the semivariogram surface in (a) for the exhaustive data verifies this choice. Generally, an exhaustive set is not available. Because the data are isotropic and assuming smoothness, only an omnidirectional experimental semivariogram is needed. An omnidirectional semivariogram is shown in Figure 3.2. For this semivariogram we have shown twelve lags with a chosen lag spacing of 2.5.

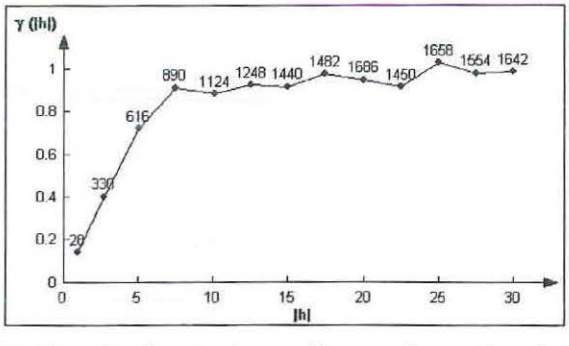

Figure 3.2: Omnidirectional semivariogram for normal scores transformed True150.

Note that the numbers shown at each point are the number of pairs at that particular lag spacing. Since the data are normal score transformed, the sill plus any nugget effect must sum to one when attempting to find a suitable model for the SGSIM algorithm. The omnidirectional semivariogram model superimposed on the experimental semivariogram can be seen in Figure 3.3. It is important to get a good fit in the early lags because it is closeness of the early lags that is mainly reflected in the simulation.

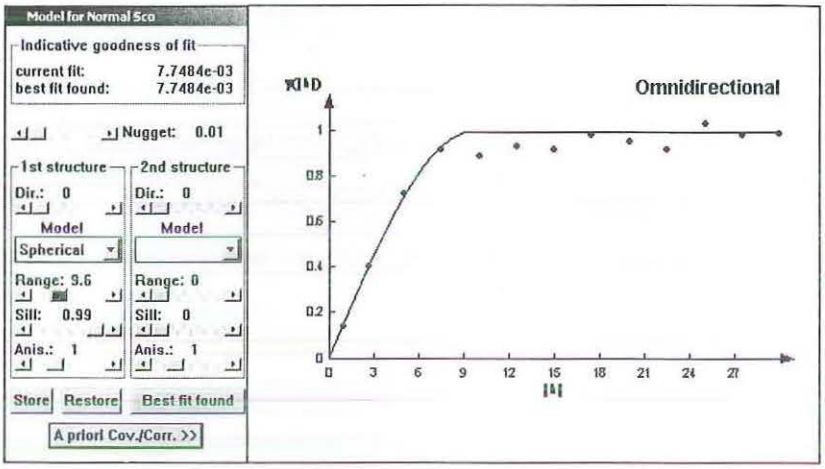

Figure 3.3: Modelled semivariogram for normal scores transformed True150.

This omnidirectional semivariogram model is a nested structure with a nugget effect of 0.01, together with one spherical structure with a range of 9.6 and a sill of 0.99. This is the semivariogram model we will use in Sequential Gaussian simulation in Section 5.2.

3.3 Variography for logarithm transformed *True150*.

Using the *True* and *True150* data that have been log-transformed, the first step is to generate semivariogram surfaces using the GSLIB program VARMAP. The semivariogram surfaces for *True* and *True150* are shown in Figure 3.4. The lag spacing for the logarithms of *True* is set at one, whilst the lag spacing for the logarithms of sample data is 2.5.

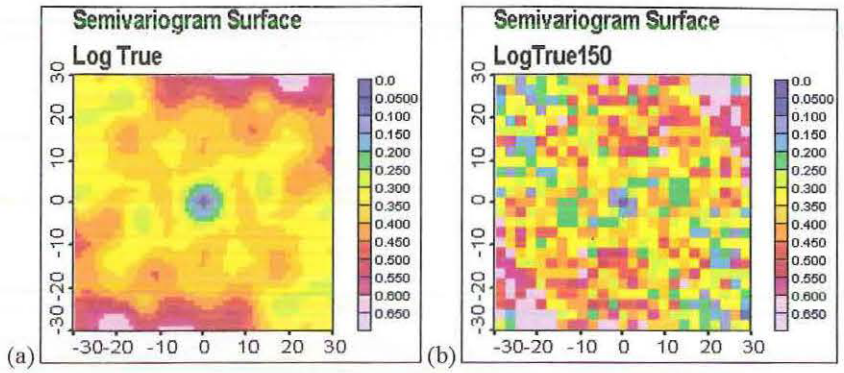

Figure 3.4: Semivariogram surfaces of logarithm data for True (a) and True150 (b).

Like the semivariogram surfaces of the normal score transformed data the semivariogram surface for the sample in Figure 3.4 shows that the data are isotropic and, once again, only an omnidirectional experimental semivariogram is needed. Also, using the semivariogram surface for the exhaustive data in (a) it is confirmed that the data are isotropic. An omnidirectional experimental semivariogram, together with the semivariogram model, is shown in Figure 3.5. It has been constructed using twelve lags with a lag spacing of 2.5.

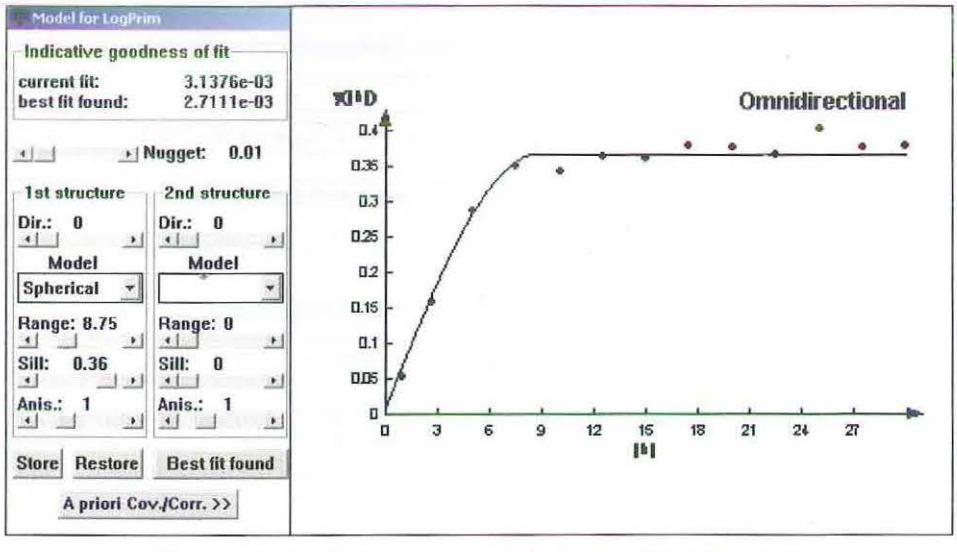

Figure 3.5: Modelled semivariogram for log transformed True150.

In Figure 3.5, we have tried to get a good fit in the early lags to the semivariogram model. This omnidirectional semivariogram model in Figure 3.5 is a nested structure with a nugget effect of 0.01, together with one spherical structure (*Sph*) with a range of 8.75 and a sill of 0.36. This is the semivariogram

model we will attempt to reproduce in Simulated Annealing using *True150* in Section 7.

3.4 Variography for raw *Berea100*.

Since *Berea100* is approximately normally distributed, the first variography is conducted using the sample data. The semivariogram surfaces for *Berea* and *Berea100* are given in Figure 3.6. The lag spacing for the exhaustive data is set at one, whilst the lag spacing for the sample data is 2.5.

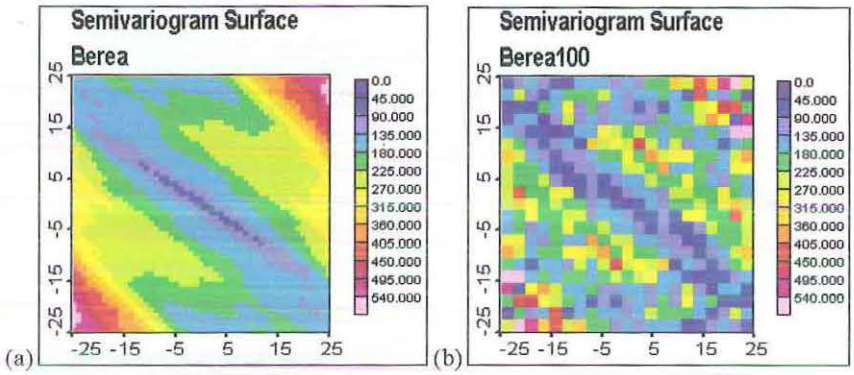

Figure 3.6: Semivariogram surfaces for Berea (a) and Berea100 (b).

From the semivariogram surface of the sample in Figure 3.6b it can be seen that the data are anisotropic and the major axis of continuity is approximately in the northwest-southeast direction. This was already indicated in Section 2.2 in our observation of the presence of banding in the northwest-southeast direction. When we compare this semivariogram surface of the sample with that of exhaustive data in Figure 3.6a, the direction of maximum continuity inferred from the experimental semivariogram surface is slightly different. Since an exhaustive data set is generally not available, the direction of maximum continuity we will use will be that from the sample of 135°. Because there is anisotropy, it is necessary to construct semivariograms at 45°, the perpendicular or minor direction and 135°, the major direction. The semivariograms in these directions are shown in Figure 3.7. The number of lags used is ten with a lag spacing of 2.5 and the angular tolerance used is 45°. Although we are using the lag spacing that is also used in *True150*, this is the smallest value we can find in which anisotropy is

clearly visible. Choosing an angular tolerance of 45° allows us to use the maximum number of points without overlapping.

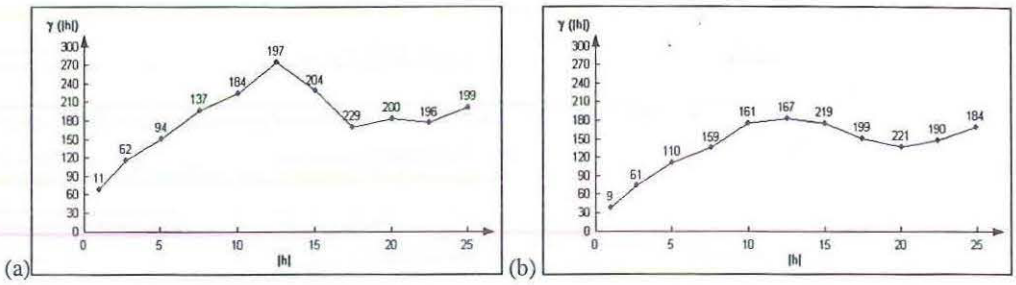

Figure 3.7: Semivariogram of raw Berea100 in directions (a) 45° and (b) 135°.

The semivariogram in direction 135° in Figure 3.7b has a much lower sill than the semivariogram in direction 45° in Figure 3.7a. The overall model is therefore one of zonal anisotropy. We shall set up this zonal model as a sum of an isotropic model and a geometric model with the large range in the direction of maximum continuity. This is seen in Figure 3.8.

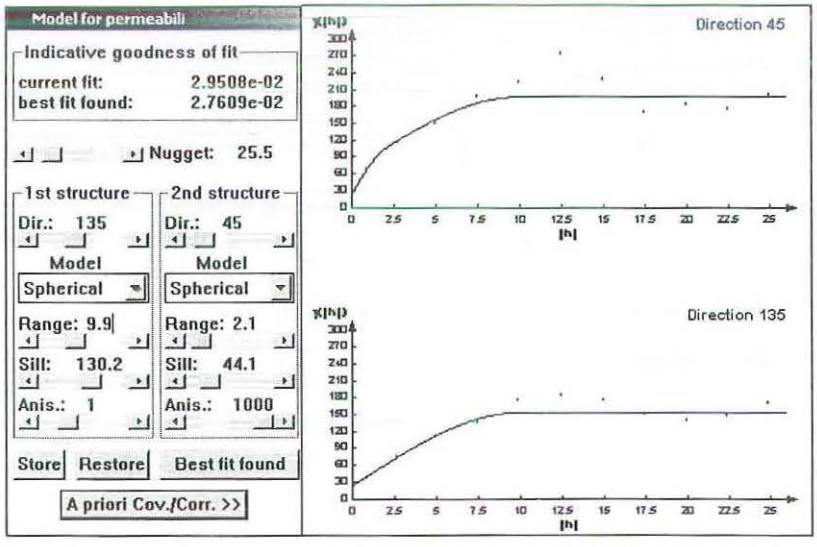

Figure 3.8: Modelled semivariogram for Berea100.

Once again we have tried to get a good fit for the early lags. The semivariogram model in Figure 3.8 consists of a nested structure with a nugget effect of 25.5, together with an isotropic spherical (*Sph*) structure with a range of 9.9 and a sill of 130.2, and a geometric spherical structure in the minor direction with a range of 2.1 and a sill of 44.1. This is the semivariogram model that is used in Simulated Annealing simulation using *Berea100* in Section 7.

3.5 Variography for NSCORE transformed *Berea100.*

Using the *Berea* and *Berea100* data that have been transformed to normal score data, the first step is to generate semivariogram surfaces using VARMAP. The resulting semivariogram surfaces for the normal scores of *Berea* and *Berea100* are shown in Figure 3.9.

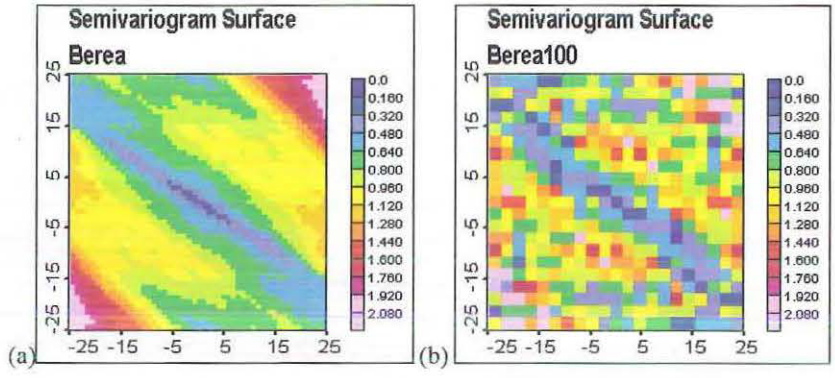

Figure 3.9: Semivariogram surfaces for Berea (a) and Berea100 (b).

As expected, the semivariogram surface in Figure 3.9b for *Berea100* show that the data are anisotropic and the major axis of continuity is approximately in the northwest-southeast direction. The semivariogram surface in Figure 3.9a further confirms that there is anisotropy, but the direction of maximum continuity is slightly different for *Berea* when compared with the semivariogram surface for *Berea100*. Since we need to make our decisions on the sample data, the experimental semivariograms in directions 45° and 135° need to be constructed. These semivariograms are shown in Figure 3.10.

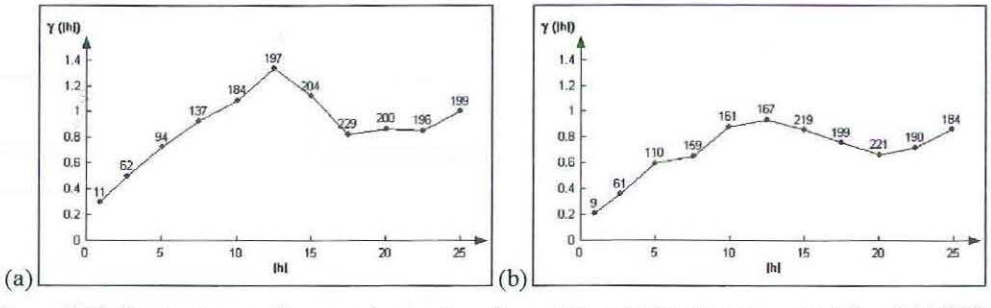

Figure 3.10: Semivariogram for normal score transformed Berea100 in directions (a) 45° and (b) 135°.

The experimental semivariograms in Figure 3.10 are similar in shape to the experimental semivariograms seen in Figure 3.7. The semivariogram in direction 135° needs to be modelled first, and then a zonal anisotropy structure is added to account for the difference in the sills. Also, the nugget effect plus the two sills for both structures need to sum to one. The modelled semivariograms are seen in Figure 3.11.

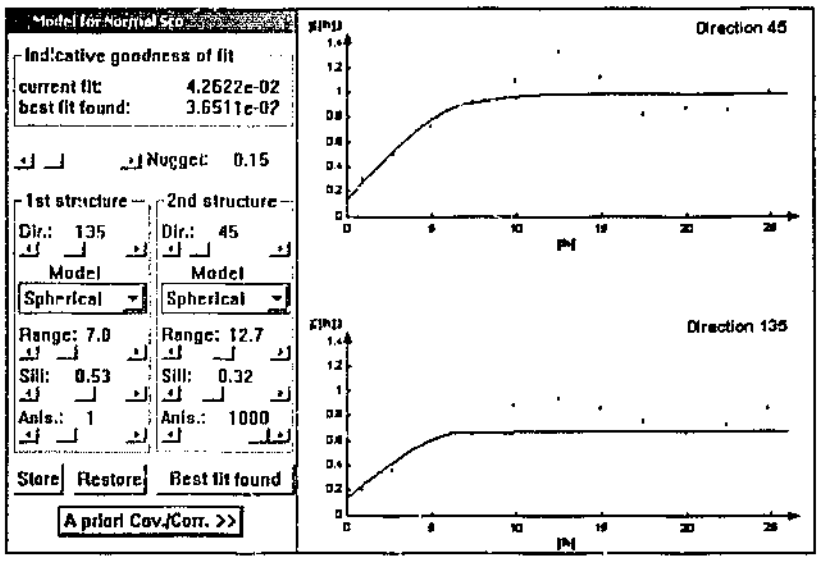

Figure 3.11: Modelled semivariogram for Berea100.

The semivariogram model in Figure 3.11 consists of a nested structure with a nugget effect of 0.15, together with an isotropic spherical structure with a range of 7.0 and a sill of 0.53 and a geometric spherical structure in the minor direction that has a range of 12.7 and a sill of 0.32. This is the semivariogram model used when we carry out Sequential Gaussian simulation using *Berea100* in Section 5.3.

4. THEORY OF SIMULATION.

4.1 Simulation

Simulation is a technique used in geostatistics to obtain values at unsampled locations. It uses the sample data set and the semivariogram model to generate a number of realizations. The simulation is called conditional if at each location the simulated value is equal to the actual value. In this thesis we concentrate on two simulation techniques, Sequential Gaussian and Simulated Annealing. Sequential Gaussian Simulation relies on simple kriging and Simulated Annealing mimics the process of annealing used in metallurgy to generate the realizations. Simulated Annealing requires objective functions and decision rules.

4.2 Simple Kriging.

One method of estimation used in geostatistics is kriging. Kriging is the term used by geostatisticians in recognition of the pioneering work associated with generalised least-squares regression algorithms done by Danie Krige (Goovaerts, 1997). It is a technique used to provide estimates for the unknown values that minimises the error variance. The kriging algorithm is a smooth interpolator in that extreme values are smoothed out, so low values are overestimated and high values are underestimated, and so the variance of the estimates is less than the original variance.

All kriging estimators are variants of the linear regression estimator $Z^*(\mathbf{u})$ given by:

$$Z^*(\mathbf{u}) - m(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha(\mathbf{u})[Z(\mathbf{u}_\alpha) - m(\mathbf{u}_\alpha)] \quad \alpha = 1,\ldots,n(\mathbf{u}) \qquad (4.1)$$

where $\lambda_\alpha(\mathbf{u})$ are the weights given to the datum $z(\mathbf{u}_\alpha)$ at location $\mathbf{u}_\alpha$ of the random variable $Z(\mathbf{u}_\alpha)$, and $m(\mathbf{u})$ and $m(\mathbf{u}_\alpha)$ are the expected values of the random variables $Z(\mathbf{u})$ and $Z(\mathbf{u}_\alpha)$ respectively.

One useful kriging method is *simple kriging* and this is used in sequential Gaussian simulation. Simple kriging, unlike other kriging methods, requires that the mean m is known and constant throughout the whole study region A. The linear estimator $Z^*_{SK}(\mathbf{u})$ for the simple kriging algorithm is obtained from (4.1) by substituting m for $m(\mathbf{u}_\alpha)$ as follows:

$$Z^*_{SK}(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{SK}(\mathbf{u}) Z(\mathbf{u}_\alpha) + \left[1 - \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{SK}(\mathbf{u})\right] m \quad \alpha = 1, \ldots, n(\mathbf{u}) \qquad (4.2)$$

where $\lambda_\alpha^{SK}(\mathbf{u})$ are the weights to minimize the error variance

$$\sigma_E^2(\mathbf{u}) = \mathrm{Var}\left\{Z^*_{SK}(\mathbf{u}) - Z(\mathbf{u})\right\} \qquad (4.3)$$

under the unbiasedness constraint $\mathrm{E}\left\{Z^*_{SK}(\mathbf{u}) - Z(\mathbf{u})\right\} = 0$. The simple kriging variance $\sigma^2_{SK}(\mathbf{u})$, using the weight $\lambda_\alpha^{SK}(\mathbf{u})$ and covariance $C(\mathbf{h})$, can be expressed as:

$$\sigma^2_{SK}(\mathbf{u}) = C(0) - \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{SK}(\mathbf{u}) C(\mathbf{u}_\alpha - \mathbf{u}) \quad \alpha = 1, \ldots, n(\mathbf{u}) \qquad (4.4)$$

4.3. Sequential Gaussian Simulation.

Sequential Gaussian simulation uses standard normal data and is based on simple kriging. The steps adopted in the Sequential Gaussian simulation of an attribute Z are:

1. If the given sample does not have a standard normal distribution, then transform the z-data into y-data with a standard normal cumulative

distribution function, using the normal score transform ϕ (introduced in Section 2.3), checking for bivariate normality of the resulting normal score variable $Y(\mathbf{u}) = \phi(Z(\mathbf{u}))$.

2. Now using the y-data:
   i. Define a random path visiting each location exactly once.
   ii. At each location $\mathbf{u}'$, find the mean and variance of the Gaussian conditional cumulative distribution function $G(\mathbf{u}'; y \mid (n))$ of the standard normal random variable $Y(\mathbf{u}')$ using simple kriging as well as the normal score semivariogram model $\gamma_Y(\mathbf{h})$. The conditioning information (n) contains a number $n(\mathbf{u}')$ of normal score data and values simulated at previously visited locations.
   iii. Draw a simulated value $y^{(l)}(\mathbf{u}')$ from the conditional cumulative distribution function and add it to the data set.
   iv. Move to the next location along the random path and repeat points (ii) and (iii).
   v. Loop until all locations are simulated.

3. The last step is then to back-transform the simulated normal scores $\{ y^{(l)}(\mathbf{u}'_j), j = 1, \ldots, N \}$ into simulated values for the original data using the inverse of the normal score transform by

$$z^{(l)}(\mathbf{u}'_j) = \phi^{-1}(y^{(l)}(\mathbf{u}'_j)), j = 1, \ldots, N. \qquad (4.5)$$

Further realizations $\{ z^{(l')}(\mathbf{u}'_j), j = 1, \ldots, N, l' \neq l \}$ are created by repeating steps two and three with a different random path.

5. USING SGSIM.

5.1 The SGSIM Parameter File.

In GSLIB the SGSIM algorithm uses an executable file together with a customised parameter file. This parameter file consists of a number of variables required to use Sequential Gaussian simulation to generate simulation realizations. Prior to creating the parameter file, it is necessary to know the following variables.

- The input data file. In this case, *True150* and *Berea100* contain the input data. When using Sequential Gaussian simulation we use the sample data sets with the assumption that the exhaustive data set is not available.
- The column number. For *True150* the column required is Primary, which is column number three. In *Berea100*, the column number is four, which is the permeability of the Berea slab.
- The file names for the debug file, the output data file, and the transformation file. The debug file will consist of the intermediate calculation steps from the running of the SGSIM program. The transformation file consists of the transformation table, which shows the normal score equivalent to the data in the input file. The output data file contains the values for each grid node, whether 2500 values for *True150* or 1600 for *Berea100*.
- Details of the cumulative frequency histogram. These histograms are seen in Figure 5.1. SGSIM needs these details to attempt to reproduce the histogram and to assist in the back-transformation procedure. The first parameters required are the minimum and maximum allowable data values. By viewing the histogram for *True150* in Figure 5.1a, we can see that the tail is still increasing and therefore the maximum allowable value should be greater than 30. The values chosen are 0, because this cannot be less than zero, and 50. For *Berea100* in Figure 5.1b the corresponding values are 20 and 95. These values were selected because the interval between the minimum and maximum values of the exhaustive and sample

data are not all that different, and our interval needs to be smaller than the interval of the exhaustive data so as to minimize outliers. The following lines in the parameter file are for the lower tail and upper tail of the cumulative frequency histogram. *True150* appears to have a linear lower tail with a hyperbolic upper tail of weight (ω) 1.25. *Berea100* has a power lower tail of weight 2.0 and a hyperbolic upper tail of weight 1.5. The power lower tail is modelled as

$$[F(z)]_{Pow.} = \left[\frac{z - z_{min}}{z_1 - z_{min}}\right]^{\omega} \cdot F^*(z_1) \quad \forall z \in (z_{min}, z_1]$$

where z_1 is the smallest z-data value and z_{min} is the minimum z-value fixed by the user. The hyperbolic tail is

$$[F(z)]_{Hyp} = 1 - \frac{\lambda}{z^{\omega}} \quad \forall z > z_k$$

where $\lambda = z_k^{\omega} \cdot [1 - F^*(z_k)]$

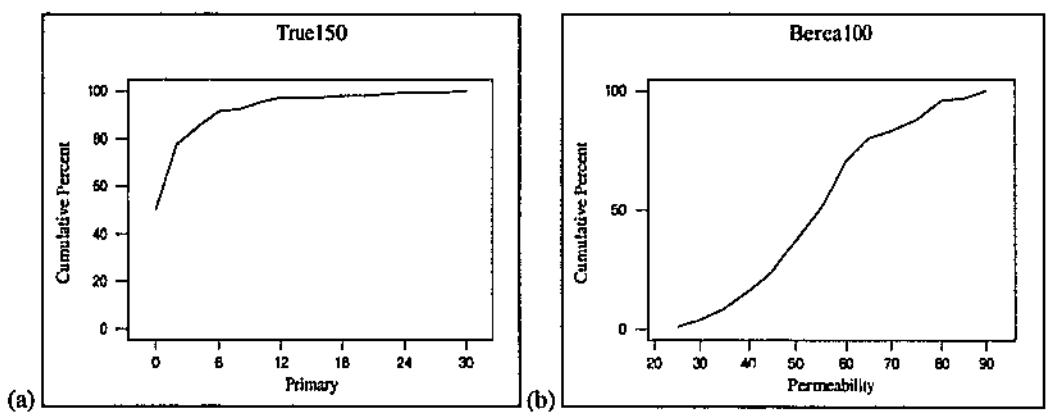

Figure 5.1: The cumulative frequency histograms for (a) True150 and (b) Berea100.

- The number of realizations to generate. We generate 50 realizations, the first two of which will be considered in detail in Section 5.2 for *True150* and Section 5.3 for *Berea100.* The first ten realizations for each sample will then be used to examine variability between these realizations in the histograms and semivariograms. All 50 realizations will then be used when we compare the 50 realizations from SASIM with these realizations in SGSIM.
- The grid size. The output grid for *True150* is 50 by 50 grid and *Berea100* it is 40 by 40.

- Minimum values for x and y, distance between grid nodes and number of nodes in each direction. Both begin at 0.5 and the grid spacing in each direction is one unit apart.
- The maximum search radius. In the case of *Berea100* where there are two structures and therefore two different ranges, the highest range is considered. For *True150*, the model is isotropic and therefore the search is circular with a radius of 10, which is selected as slightly larger than the range of the semivariogram model. In the case of *Berea100* where the model is anisotropic, there are two different ranges. The higher range is considered. Therefore we search using an ellipse at angle of 135°, with a maximum radius of 30 and a minimum of 15.
- The semivariogram model, which was examined in section 3.2 for *True150* and section 3.5 for *Berea100.*

The resulting parameter files for Sequential Gaussian simulation of *True150* and *Berea100*, are given in Appendix C.5. Once these parameter files are created, they are run using the executable file SGSIM, which in turn generates 50 realizations for each data set.

5.2 Sequential Gaussian Simulation using *True150.*

After running SGSIM for *True150*, the output file contains values for 50 realizations, each with 2500 values assigned on a 50 by 50 grid. But all the data are found in one column for the 50 realizations and need to be separated. We copied this column into Excel and divided it into 50 columns with 2500 values. Then two columns were copied to Minitab where we created a new file that included two extra columns for the coordinates. Each column for the realizations was then mapped using 3Plot. The resulting maps of the first two realizations are shown in Figure 5.2, together with the exhaustive and sample data maps for comparison.

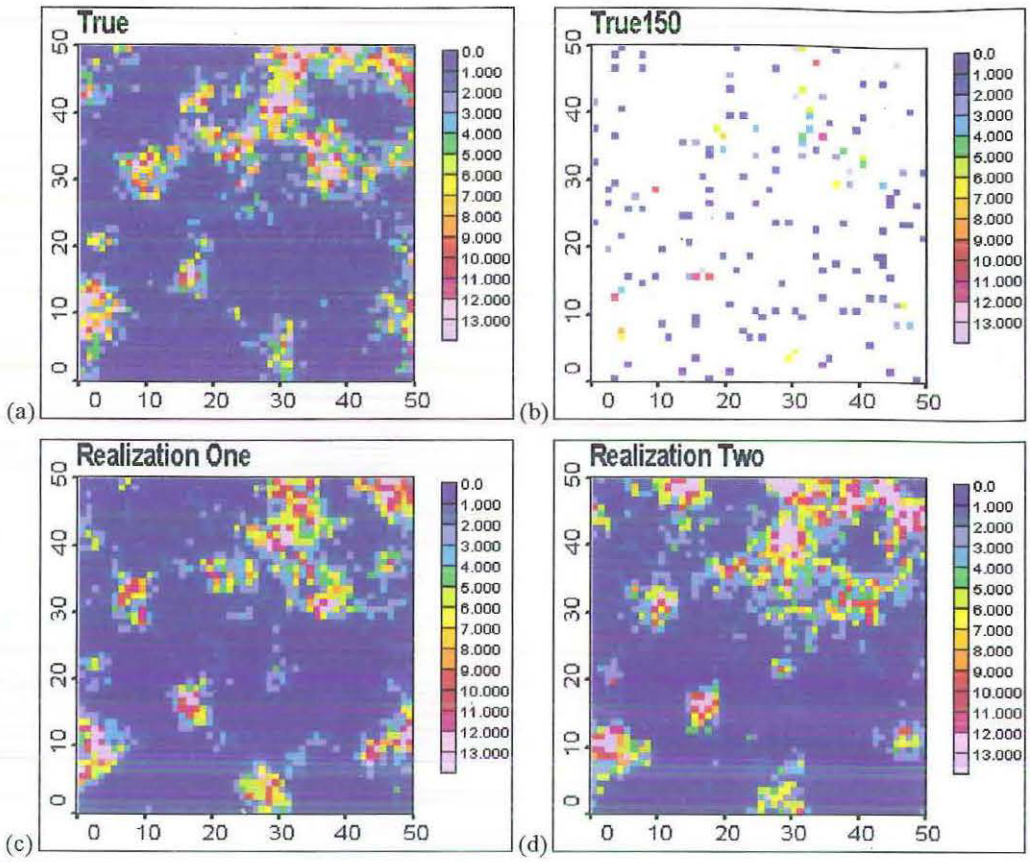

Figure 5.2: The exhaustive and sample maps (a) and (b) and two SGSIM realization maps (c) and (d).

Since we have a map of the exhaustive set in (a), we can use it to compare with the output from the two realizations in Figure 5.2. It seems when we compare the realizations they are similar. If we compare the realizations with the sample data map in (b) we see that the values of the locations are the same in both. Also, the regions of high values in the sample data have been reflected in the maps of the realizations, especially in the top right hand corner.

The omnidirectional semivariograms are now needed to see how well the semivariogram model has been reproduced. Before generating the omnidirectional semivariograms, it is necessary to transform the data into standard normal data using NSCORE, since the semivariogram model is for standard normal data. Once the data are transformed, the GSLIB program GAMV is used to generate the semivariograms. The output from GAMV is then copied to Excel and the omnidirectional semivariograms created. These are shown in Figure 5.3, together with the model to be reproduced, and use 15 lags with a lag spacing of one.

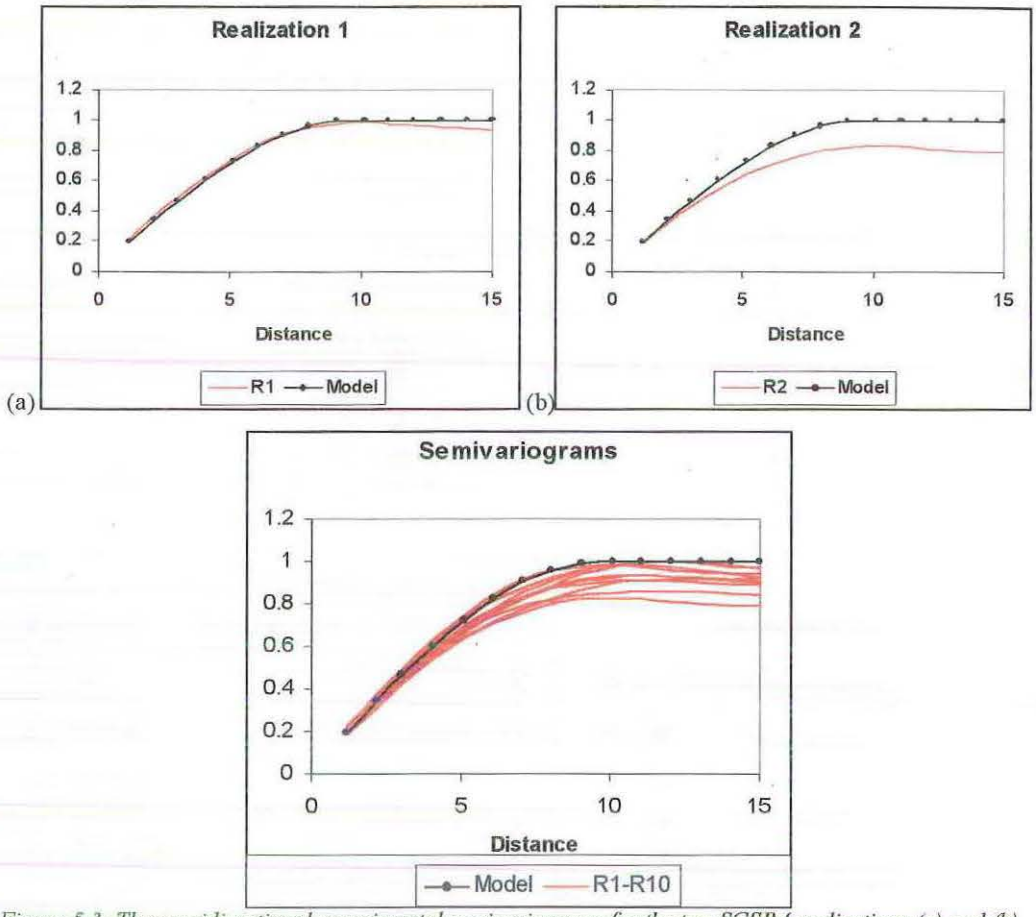

Figure 5.3: The omnidirectional experimental semivariograms for the two SGSIM realizations (a) and (b), and for 10 realizations (c). Each includes the semivariogram model.

In Figure 5.3, we see the experimental semivariograms of the two realizations in (a) and (b) together with the required model. When compared to the required model, we see that the semivariogram of the first realization closely matches the model, but the semivariogram of the second realization does not. This shows that not all realizations are going to reproduce the model exactly. To demonstrate this we have in (c) omnidirectional experimental semivariograms for the first ten realizations, together with the required model. Out of the ten only three realizations closely match the semivariogram model, while other realizations underestimated it. As can be seen in Figure 5.3 the experimental semivariogram of most realizations do not reach the sill, however most approximate the range of the semivariogram model. A possible explanation for this feature is the manner in which the normal score for the semivariogram calculations were reproduced. The normal scores were generated from the back transformed simulated values and can potentially differ from the normal scores generated through SGSIM.

Since the SGSIM parameter file includes the cumulative frequency histogram, it may be useful to view the histograms of the realizations to see if histogram of the sample is reproduced. The histograms for each realization and for *True150* are in Figure 5.4.

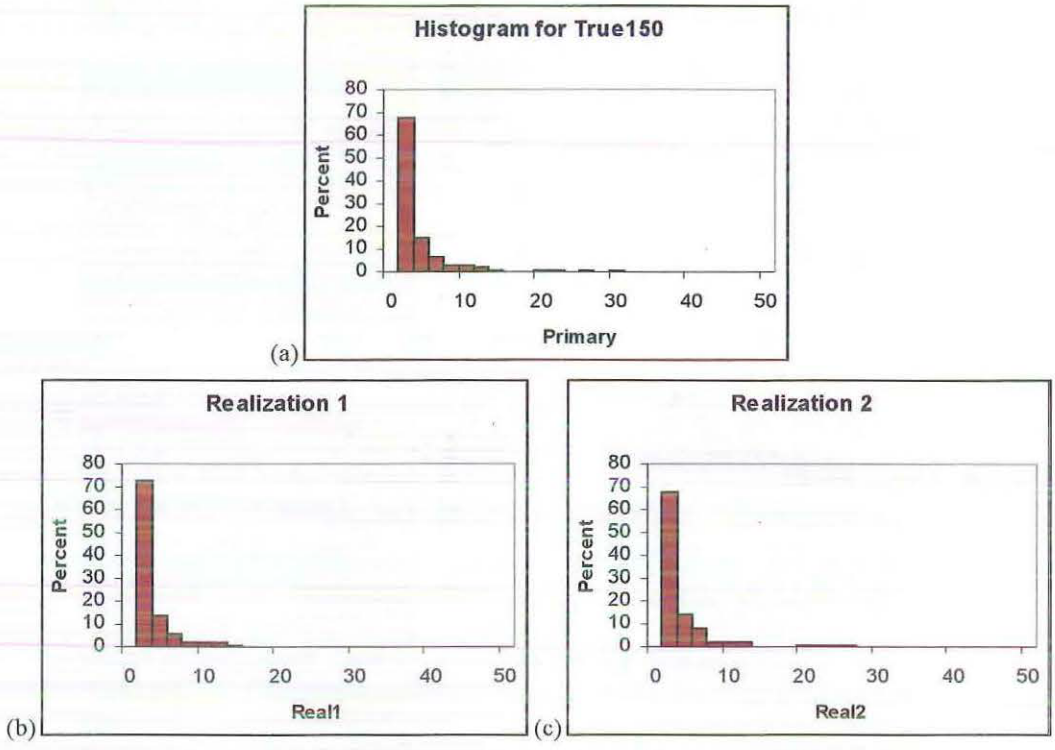

Figure 5.4: The histograms for True150 (a) and the realizations (b) and (c).

The histograms of the realizations in Figure 5.4 all begin with a high number of values less than two, then taper off to about value twelve, and contain several outliers to 50. The only difference between the histograms for the realizations and the histogram of the sample in (a) is the range. The range for the histograms of the realizations is 50 while the histogram of the sample is approximately 30. Furthermore, because we are using percentages, we can see that there is approximately 68 percent of values between zero and two in the sample, and around 70 percent in the realizations.

5.3 Sequential Gaussian Simulation using *Berea100.*

Using the same procedure as in section 5.2, where we created 50 realizations and used Excel and Minitab to create a new data file, we have two

columns of data, each with 1600 values. Each realization is then mapped using 3Plot. The resulting maps of the realizations are in Figure 5.5, together with the original exhaustive and sample data map for comparison.

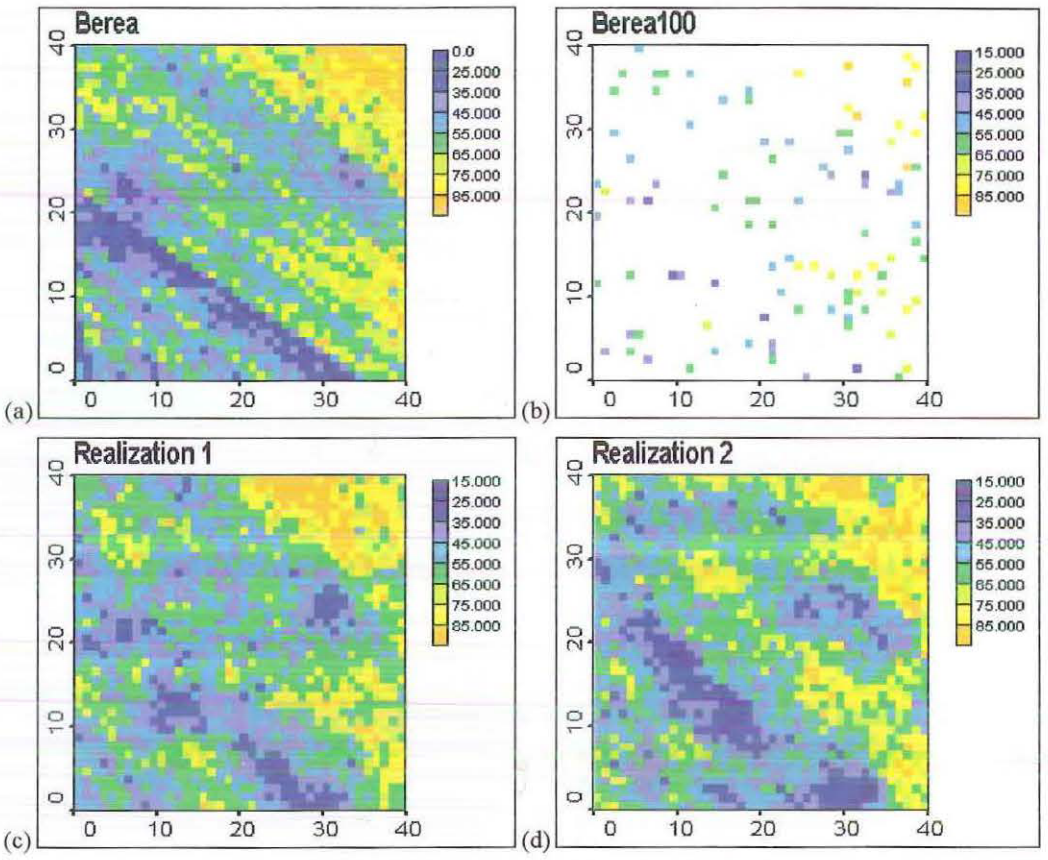

Figure 5.5: The exhaustive and sample data maps (a) and (b), and two SGSIM realization maps (c) and (d).

The two maps of the realizations in Figure 5.5 when compared with the exhaustive data map in (a) are similar in that the high values are found in the top right-hand corner, and the low values are seen in a band. In the map of the sample we see that there are regions where there are low and high values. These regions are reflected in the maps of the realizations.

As done with the *True150* realizations, these realizations need to be transformed into standard normal data using NSCORE. Once transformed, the semivariograms in directions 45° and 135° can be generated using GAM. The semivariograms for the realizations are in Figure 5.6. The semivariograms use 14 lags with a lag spacing of $\sqrt{2}$.

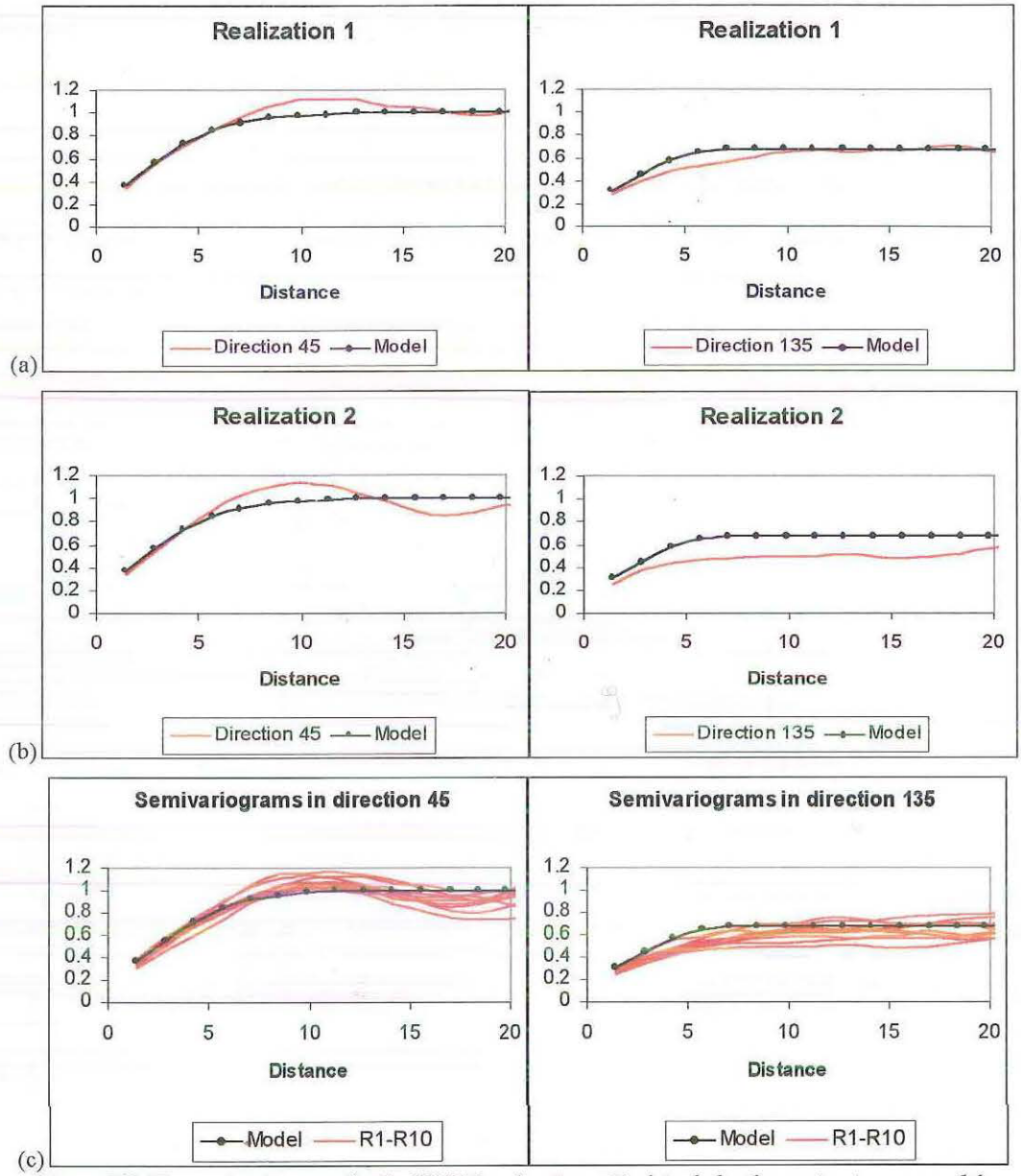

Figure 5.6: The semivariograms for the SGSIM realizations. Each includes the semivariogram model.

The semivariograms shown in Figure 5.6 for the realizations are in the directions 45° and 135°. In (a) and (b), we see the semivariograms of the two individual realizations. In (c), the semivariograms shown are for the first ten realizations in directions of maximum and minimum continuity. In all cases the required semivariogram models are included as a black line, whilst the semivariograms for the realizations are in red. The semivariograms in direction 45° for the realizations tend to follow a similar pattern. That is, they follow the initial lags, then peak above the model, and lastly fall under the model. For the semivariograms in direction 135°, we see that some are below the model while some others are above the model, but the first few lags for all realizations are

below the value of the lag in the model. The better reproduction of the semivariogram model is from the first realization.

Next we need to check if the sample histogram has been reproduced in the histograms for the realizations. The histograms for each realization and for *Berea100* are in Figure 5.7.

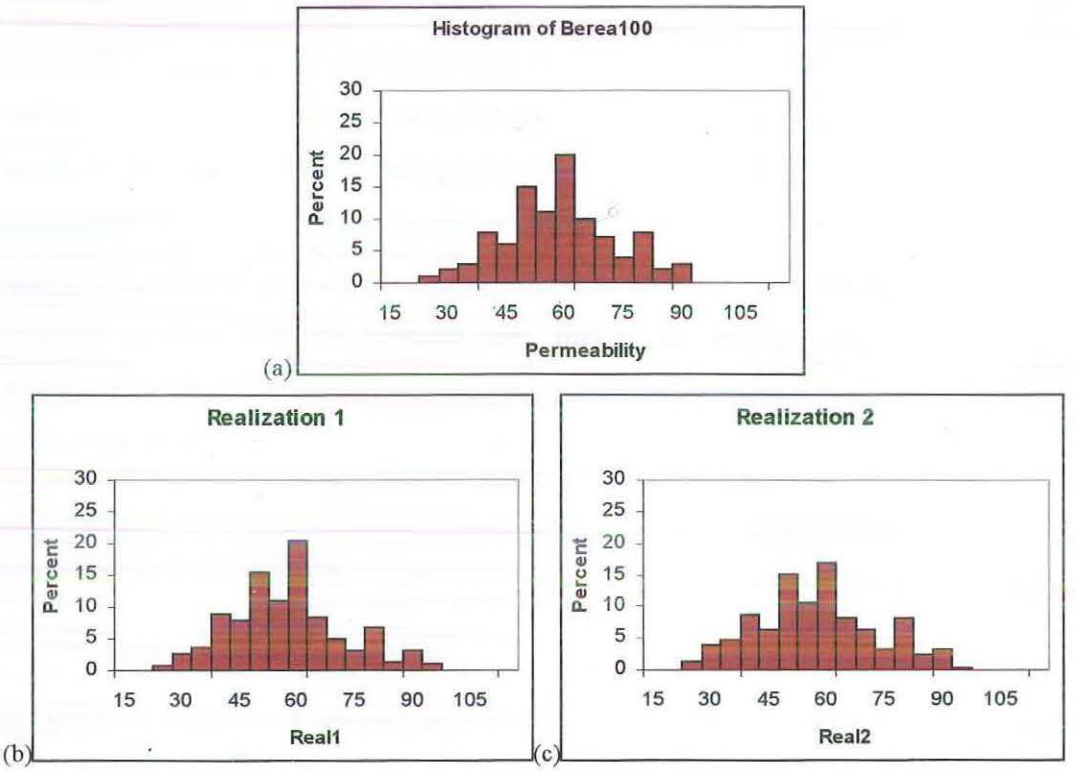

Figure 5.7: Histograms for theBerea100 (a) and the realizations (b) and (c).

The histograms for the realizations in Figure 5.7 show that, when compared to the histogram for the sample data in (a), they contain similar aspects. They all appear to have normally distributed data, high number of values in the centre, with low number of values tapered off at the ends. The high values at classes 50 and 60 in the histogram of the sample have been reproduced in the histograms of the first two realizations.

6. SIMULATED ANNEALING.

6.1 Annealing.

The term annealing comes from a physical process developed to increase the workability and durability of materials such as metals, alloys and glass. Annealing of materials begins by heating the material to a high temperature just under melting point, and slowly cooling the temperature, such that enough time is spent at each temperature to reach thermal equilibrium so as to permit a reordering of the molecules to a highly structured state. If the time spent at each temperature is insufficient then the probability of getting a very low energy configuration, especially around freezing point, is greatly reduced. This slow heating and cooling relieves internal stresses and makes up dislocations or vacancies introduced during the mechanical shaping. This annealing process is the basis for simulated annealing.

6.2 Simulated Annealing.

The concept of Simulated Annealing was first introduced in Metropolis et al (1953) as a modified Monte Carlo scheme, but it was not until Kirkpatrick et al (1983) that it was termed simulated annealing. Metropolis et al used the concept to calculate the properties of different substances composed of interacting molecules. Kirkpatrick et al introduced it as a technique to solve a number of problems arising in the design of computers and in the study of the travelling salesman problem consisting of several thousand cities. Simulated Annealing is the generic name given to a set of mathematical non-linear optimisation algorithms that are used to solve large-scale problems.

Simulated Annealing is used in geostatistics to obtain different realizations that have the same spatial characteristics. Compared with other simulation techniques, Simulated Annealing is a relatively new approach that generates alternate conditional images of either continuous or categorical variables and has the potential to combine two-point statistics - when variables relate to the same

attribute at two locations - with complex multiple-point spatial statistics (Deutsch and Journel, 1998).

In this simulation approach there is minimal reference to a random function model. Rather, simulated annealing is formulated as an optimisation problem to be solved by a numerical optimisation technique. For this the requirements for simulated annealing are:

- An initial image
- A perturbation mechanism
- An objective function
- A decision rule
- A convergence criterion

In geostatistics, using the continuous attribute z at N locations, Simulated Annealing requires that an objective function be defined that measures the deviation between the target and the current statistics of the realization at each perturbation. The process of Simulated Annealing is detailed below.

1. Create an initial realization $\{ z_{(0)}^{(l)}(\mathbf{u}'_j), j = 1,\ldots,N \}$ that reproduces the data values at their locations. This realization may approximate some of the target statistics such as variance and sill of the target z-semivariogram.

2. Compute the initial value of the objective function O for the initial realization.

3. Perturb the realization by selecting a single location $\mathbf{u}'_j$ at random and modifying the corresponding z-value $z_{(l-1)}(\mathbf{u}'_j)$ according to some prespecified mechanism such as resampling the target histogram with a different value $z_{(l)}(\mathbf{u}'_j) = F^{-1}(p_j)$ where (p_j) is a random number in [0,1] (Goovaerts, 1997).

4. Determine the effect of the perturbation on the reproduction of the target statistics by recomputing the objective function O_{new} accounting for the changes of the previous image.

5. Accept or reject the perturbation based on a decision rule. The decision rule used is to accept unfavourable perturbations based on the Boltzmann distribution. That is,

$$\text{Probability}\{\text{accept the i}^{\text{th}}\text{ perturbation}\} = \begin{cases} 1 & \text{if } O_{new} \le O_{old} \\ \exp\left(\dfrac{-\Delta O}{t}\right) & \text{otherwise} \end{cases}$$

where $\Delta O = O_{new} - O_{old}$, and t is the temperature. If the temperature is large, then the probability of accepting the ith unfavourable perturbation is higher.

6. If accepted, update the initial realization into a new image $\{ z^{(l)}_{(1)}(\mathbf{u}'_j), j = 1, \ldots, N \}$ with the objective function value now $O = O_{new}$.

7. Repeat steps three to six until the target constraints are met or the perturbations do not effectively reduce the objective function.

Further realizations $\{ z^{(l')}(\mathbf{u}'_j), j = 1, \ldots, N \}$ are generated by following the process above starting from different initial images.

The Simulated Annealing process appears to be inefficient in that up to a million perturbations may be required to obtain an image that contains the prespecified spatial structure (Deutsch and Cockerham, 1994). Therefore it is necessary to appropriately code the objective function in order to reduce the computer processing time. For example, when a perturbation is being considered, SASIM updates the semivariogram lags rather than recalculating, so that if a value z is a lag distance $\mathbf{h}$ away from z_i, then the previous contribution of z_i is deducted

from the semivariogram and the new contribution of z_j is then added (Deutsch and Journel, 1998).

6.3 The Initial Image.

Before Simulated Annealing can be conducted, an image must be generated. The prequisites of the initial image are:

- it needs to be easily generated,
- it should basically match the target constraints, for example, it must honour the data or reproduce the histogram, and
- all initial images are equally probable in that each is equally likely to be drawn (Goovaerts, 1997).

It is generated by freezing data values at their locations and assigning a z-value drawn at random from the target cumulative distribution function F(z) to the unsampled locations. This approach is quick and yields a set of initial images that honour the conditioning data and approximate the target histogram (Goovaerts, 1997).

6.4 The Objective Function.

Simulated Annealing requires an objective function to be declared to measure the deviation between the desired features and those of the realization. The objective function is used to create Simulated Annealing realizations that replicate the histogram or semivariogram from the original data. Generally an objective function O is made up of the weighted sum of C components:

$$O = \sum_{c=1}^{C} \omega_c O_c \qquad (6.1)$$

where ω_c is the weight of the component objective function O_c. The purpose of this weight is to control the contribution of each component to the overall objective function.

Deutsch and Journel (1998) discuss five components that could be used in the objective function used in the parameter file for the GSLIB program SASIM. For this thesis, we have decided to use two of these. These components we use are:

A *histogram*, that should be reproduced by all realizations. The cumulative distribution function $F^*(z)$ of the realization should match the cumulative distribution function F(z) of the initial data for some number of z-values. The objective function for the histogram is written as:

$$O_1 = \sum_z \left[F^*(z) - F(z)\right]^2 \qquad (6.2)$$

A *semivariogram*, which captures the two-point spatial variability in the realization. The semivariogram $\gamma^*(\mathbf{h})$ for the realization should match the prespecified semivariogram model $\gamma(\mathbf{h})$. This deviation is measured as:

$$O_2 = \sum_{\mathbf{h}} \frac{\left[\gamma^*(\mathbf{h}) - \gamma(\mathbf{h})\right]^2}{\gamma(\mathbf{h})^2} \qquad (6.3)$$

The division by the square of the $\gamma(\mathbf{h})$ value is used to standardize the units and give more weight to the closely spaced values.

6.5 The Annealing Schedule.

The annealing schedule parameters need to be selected so that the temperature is reduced effectively, such that the timing and magnitude of the reduction is sufficient to generate appropriate realizations. In the annealing schedule, the following parameters are used:

- t_0 : the initial temperature
- λ : the reduction factor, $0 < \lambda < 1$

- K_{max} : the maximum number of attempted perturbations at any one temperature. Whenever K_{max} is reached, the temperature is reduced by the factor λ.
- K_{accept} : the number of perturbations to accept. Once K_{accept} perturbations are accepted, the temperature is reduced by factor λ.
- O_{min} : the minimum number (usually 0.001) to stop the annealing schedule when the temperature is reduced by λ.
- S : the stopping number. Once K_{max} is attained S times, then the algorithm is stopped.
- ΔO : the difference between the result of the new objective function and the previous result.

The annealing schedule starts with a high temperature t_0 that permits a large number of unfavourable perturbations to be accepted at the beginning of a simulation. Throughout the simulation, the temperature is reduced by a factor of λ so as to limit the discontinuous modifications of the image. This reduction occurs whenever enough perturbations are accepted, K_{accept}, or a specified number have been tried, K_{max}. The annealing schedule ceases when the low value O_{min} is reached or when K_{max} has been reached S times.

6.6 The Convergence Criterion.

For the Simulated Annealing algorithm, like other optimisation algorithms, it is important that the process has a stopping point. The process terminates if the objective function reaches the predetermined minimum value of O_{min}, or if the maximum number of perturbations at the same temperature is exceeded, or if the number of accepted perturbations is smaller than a given threshold value (Goovaerts, 1997). All of these stopping methods are used in SASIM.

6.7 Flowchart of Simulated Annealing.

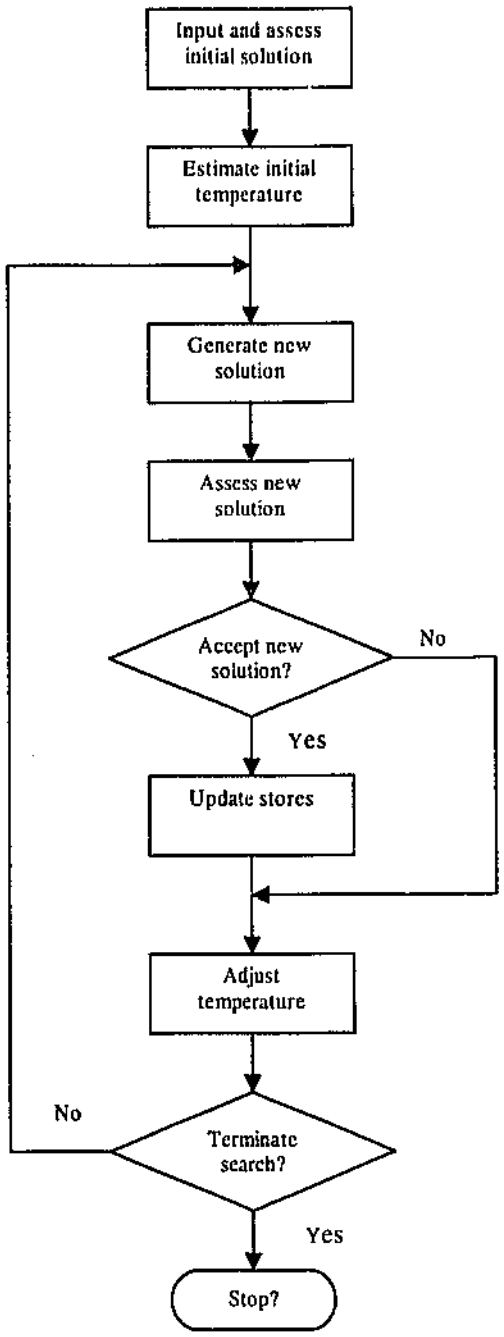

Figure 6.1: Flowchart of Simulated Annealing.

The flowchart in Figure 6.1 was reproduced from http://csep1.phy.ornl.gov/mo/node28.html. In our case we should replace "Update stores" with "Update scores" and "Stop?" with "Stop" since this is not a yes/no choice.

7. USING SASIM.

7.1 The SASIM Parameter File.

The SASIM algorithm in GSLIB is used for Simulated Annealing, and the parameter files used in this thesis are located in Appendix C.6 for *True150* and *Berea100*. The choices to be made in the SASIM parameter file are:

- The objective function components (line 1 in parameter file) and their weights (line 2) that are used. These are described in Section 6.5. The first parameter file (*TrV.par* for *True150* and *BerV.par* for *Berea100*) uses the semivariogram as the only component included in the objective function (Section 7.3). In the second parameter file (*TrVH.par* for *True150* and *BerVH.par* for *Berea100*), both a histogram and a semivariogram are included, with the weight for the semivariogram initially set to five (Section 7.4). Also, testing to see if setting the weights of the semivariogram to one, three and five, makes a difference, will be studied (section 7.5).
- Whether a transformation of the data is required. The transformation used in SASIM is logarithm to base ten. A logarithmic transformation converts positively skewed data to approximately normally distributed data, assuming that the original data are approximately lognormal. The SASIM program converts the data to logarithmic data prior to simulation and converts the data back prior to concluding by using ten to the power of the logarithmic data ($z = 10^y$, where y is the logarithmic data, and z is the original data equivalent). For *True150*, this parameter will be set to transform the data, while *Berea100* will not be transformed.
- The number of realizations. This parameter is set to 50 for each parameter file, but only two will be shown when comparing with other SASIM realizations. The 50 realizations will be used when we compare simulation methods.
- The grid definition. Looking at the mosaic map of the exhaustive data, we can see that a 50 by 50 grid is used in *True*, and a 40 by 40 grid is used in

Berea, with grid locations at a distance of one unit apart, starting at x and y equal to 0.5, and finishing at x and y equal to 49.5 for *True* and 39.5 for *Berea*. Because the data are two-dimensional, the number of z-values is set to one.

- The debug and output data files.
- Whether to use SASIM's default annealing schedule. Setting the annealing schedule to 0 uses the automatic default values. SASIM's default values are:
  - The initial temperature is 1.0.
  - The reduction factor is 0.1.
  - The maximum number of attempted perturbations is 75.
  - The number of perturbations to accept is 8.
  - The stopping number is 3.
  - The minimum value is 0.001.

  Other recommended annealing schedules are shown in Table 7.1. These annealing schedules will be studied in section 7.6.
- The input data files to be used. We use *True150* and *Berea100*. The columns for x and y are in columns one and two. The attribute Primary in *True150* is in column three, and the attribute Permeability in *Berea100* is in column four.
- The details of the histogram. The files are the input data files and the column is the attribute column for each. The number of quantiles is also required. This parameter will be set to 48 for *True150* and 24 for *Berea100*.
- The details of the indicator semivariograms. The number of indicator semivariograms and the thresholds to be used. Due to time constraints, indicator semivariograms were not considered in this thesis.
- The details of the secondary data. This is not applicable here.
- The number of lags to use for the semivariograms. Section 7.2 contains further details and the experimentation results.
- The details of the semivariogram and indicator semivariogram models. These details are the number of structures and the nugget effect, the structure type and sill, and the range found. These details are found in

Sections 3.3 and 3.4. These are the semivariogram models that we wish to reproduce in the resulting realizations.

7.2. Setting the number of lags in SASIM.

Setting the number of lags in SASIM is different from setting the number of lags we saw in the variography section. So for clarity, we shall refer to the number of lags in SASIM as *nlags*, as it is actually called in the SASIM parameter file. It is the number of lags that need to be matched in the SASIM algorithm.

This parameter needs to be selected so that the target semivariogram is appropriately reproduced. In the isotropic case, nlags should be selected so that the pairs (x,y) lie within a circle of radius less than or equal to the range of the semivariogram model.

The grid locations that are used for this reproduction corresponding to different nlags values can be seen in Figure 7.1 for *True150*. The values for nlags considered in the graph for *True150* in Figure 7.1 are 40, 56, 74, 98, and 120. The result for nlags greater than 120 has also been included.

The values 40, 56, 74 and 98 are selected because they correspond to the number of integer nodes that lie within a circle of radius five, six, seven and eight, respectively. The value 120 is selected as the maximum, because this is the number of grid locations with a distance below the range of the proposed semivariogram model. That is, when the value for nlags is entered, SASIM chooses that number of grid locations, selecting the closest grid location not yet selected. If the distance from the origin to the grid location is greater than the range, then SASIM starts choosing grid locations along the *y*-axis. The inclusion of the value greater than 120 is to demonstrate what happens when we choose an nlags value that causes the program to look outside the range. The range of the semivariogram model in the *True150* parameter files is 8.75 so the distance from the origin to each grid point must be less than this value. Because the grid is two-dimensional and the minimum and maximum search radii are equal, the distance is calculated by using the formula:

Length from origin to grid point = $\sqrt{(x^2 + y^2)}$.

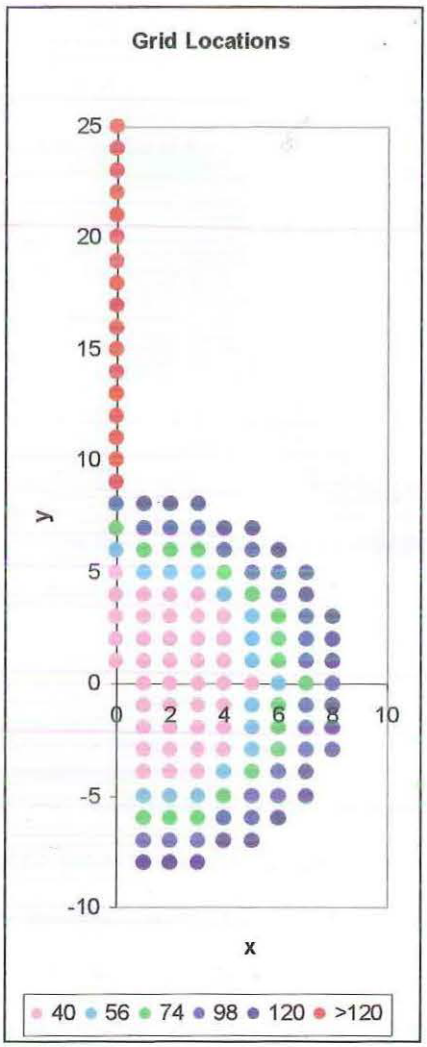

Figure 7.1: Grid locations for different nlags for True150.

To reproduce the semivariograms, the grid locations corresponding to different nlags values are shown in Figure 7.2 for *Berea100*. The nlags values for *Berea100* in Figure 7.2 are 80, 153 and 216. The result for nlags greater than 216 has also been included.

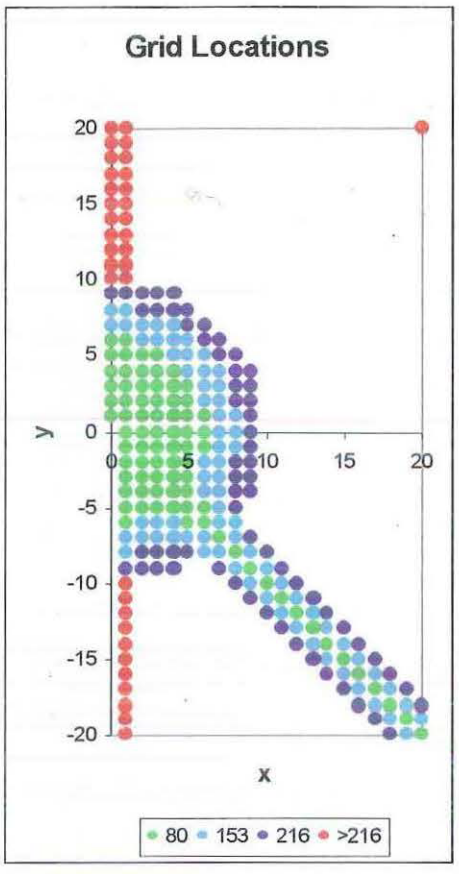

Figure 7.2: Grid locations for different nlags for Berea100.

The nlags values 80, 153 and 216 are selected because they produce circles of radii six, eight and nine, respectively, plus an extension at 135°, which is the anisotropy seen for *Berea100*. When nlags becomes greater than 216, SASIM starts selecting locations that have a distance greater than the largest range of 9.9 seen in the proposed semivariogram model. This is demonstrated by the inclusion of the nlags value greater than 216 in Figure 7.2.

For this thesis when testing the results from different nlags values, the *True150* nlags parameter we will use is the small value of 56, the middle value of 98, and the large value of 120 and for *Berea100* the nlags parameter used will be 80, 153 and 216.

7.3 Using only the semivariogram component in SASIM.

True150.

The SASIM parameter file *TrV.par* in Appendix C.6 has only the semivariogram component parameter set, but the nlags is changed. The three parameter files with the nlags equal to 56, 98 and 120, were run separately in the SASIM program. We then used Minitab to transform the data for each realization to logarithmic data. The data for each realization were used to generate realization maps for comparison with both the sample and the exhaustive data. The maps of the first two realizations for each nlags are in Figure 7.3, together with the exhaustive and sample data maps.

For each map of the realizations in Figure 7.3, it can be seen that there are some sections which are similar to the map for the exhaustive data in (a), but there appears to be one region in all realization maps at approximately (30, 20) that consist of high values not seen in the exhaustive data map. The observations from the sample map give us some idea to where there are regions of high values. These regions of high values are seen in the realizations.

The maps of the realizations in Figure 7.3 all use data that have been back transformed from SASIM. But because the semivariograms and semivariogram surfaces of the sample data in Section 3 all use log-transformed data, it is necessary to transform the data from the realizations so as to compare the semivariogram surfaces and semivariograms of the realizations. Using the parameter file VARMAP in GSLIB we can create semivariogram surfaces of the logarithmic data. The semivariogram surfaces using the logarithmic data are shown in Figure 7.4, and all show 30 lags with a lag spacing of one.

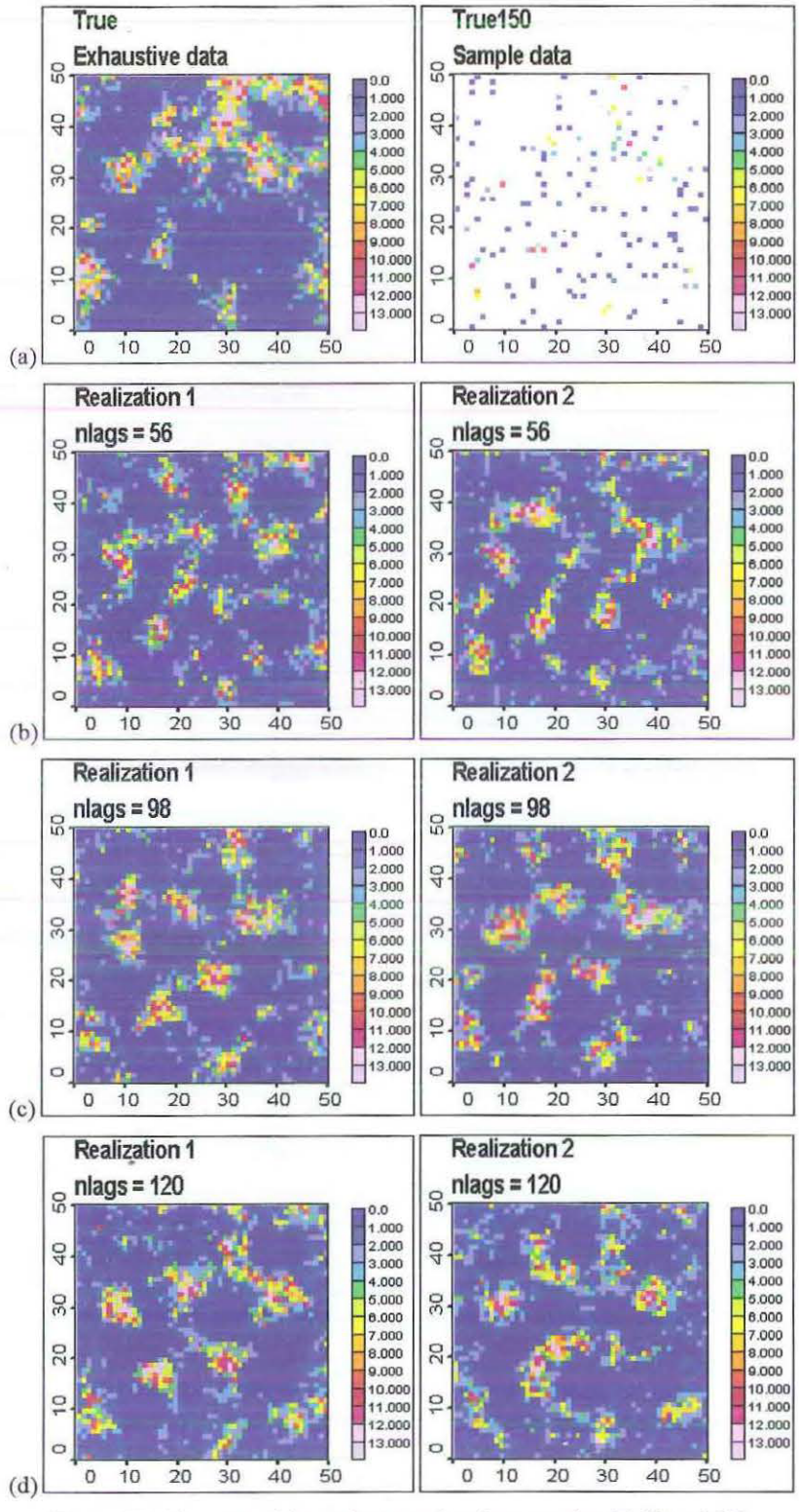

Figure 7.3: The maps of the realizations for nlags equal to 56, 98 and 120.

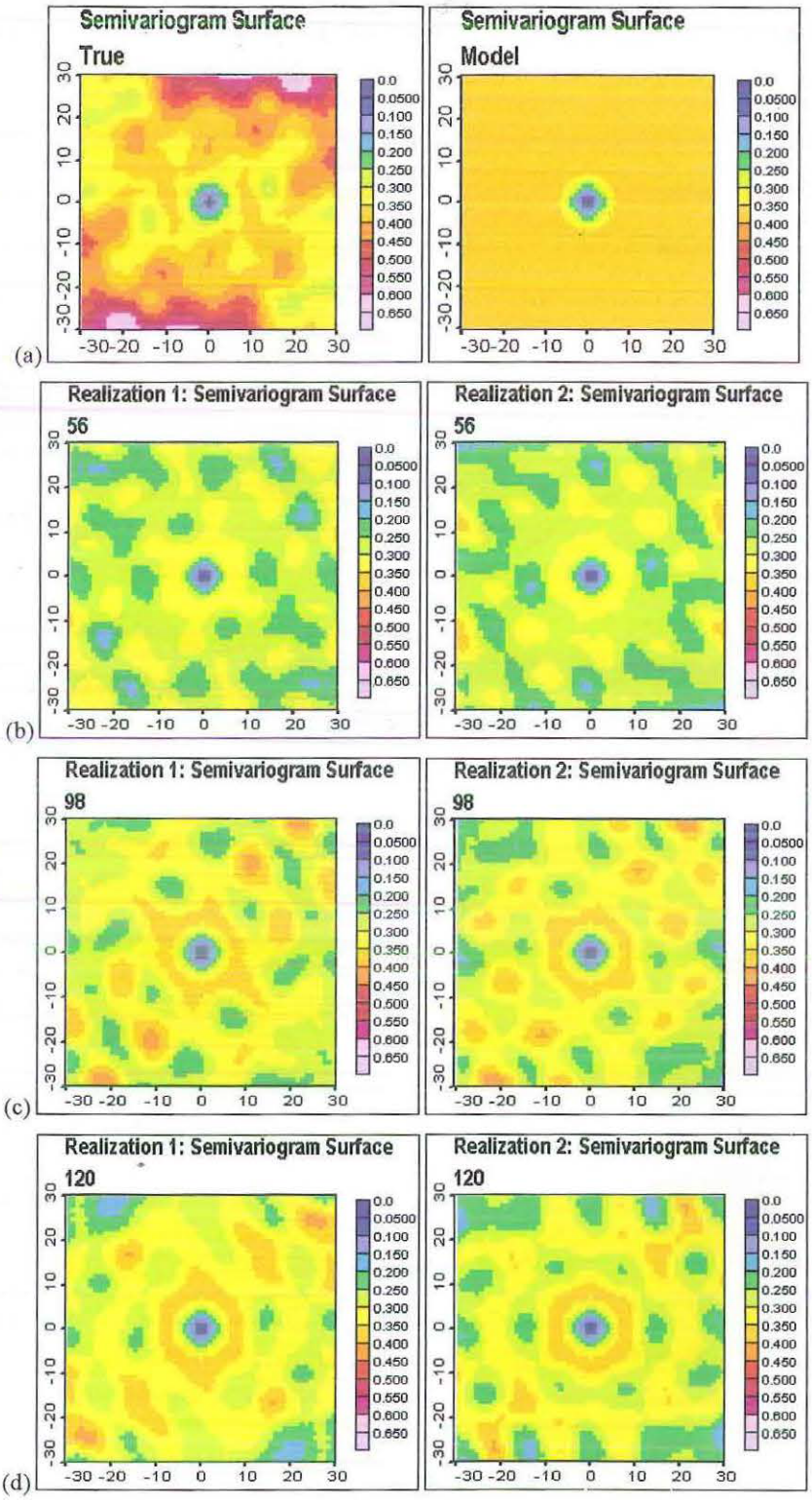

Figure 7.4: Semivariogram surfaces for the realizations for nlags equal to 56, 98 and 120.

From the semivariogram surfaces for the realizations shown in Figures 7.4, we can see that the isotropy seen in the exhaustive data map shown in (a) has been maintained. Furthermore, the semivariogram surfaces contain lower values than seen in the semivariogram surface for the exhaustive set in (a). In Figure 7.4b, we

have included a semivariogram surface for the actual model. If we compare this with the realizations, we can see that the small lags are identical.

The next step is to generate omnidirectional experimental semivariograms for each realization. These can be seen in Figure 7.5 for nlags equal to 56 (a), 98 (b) and 120 (c).

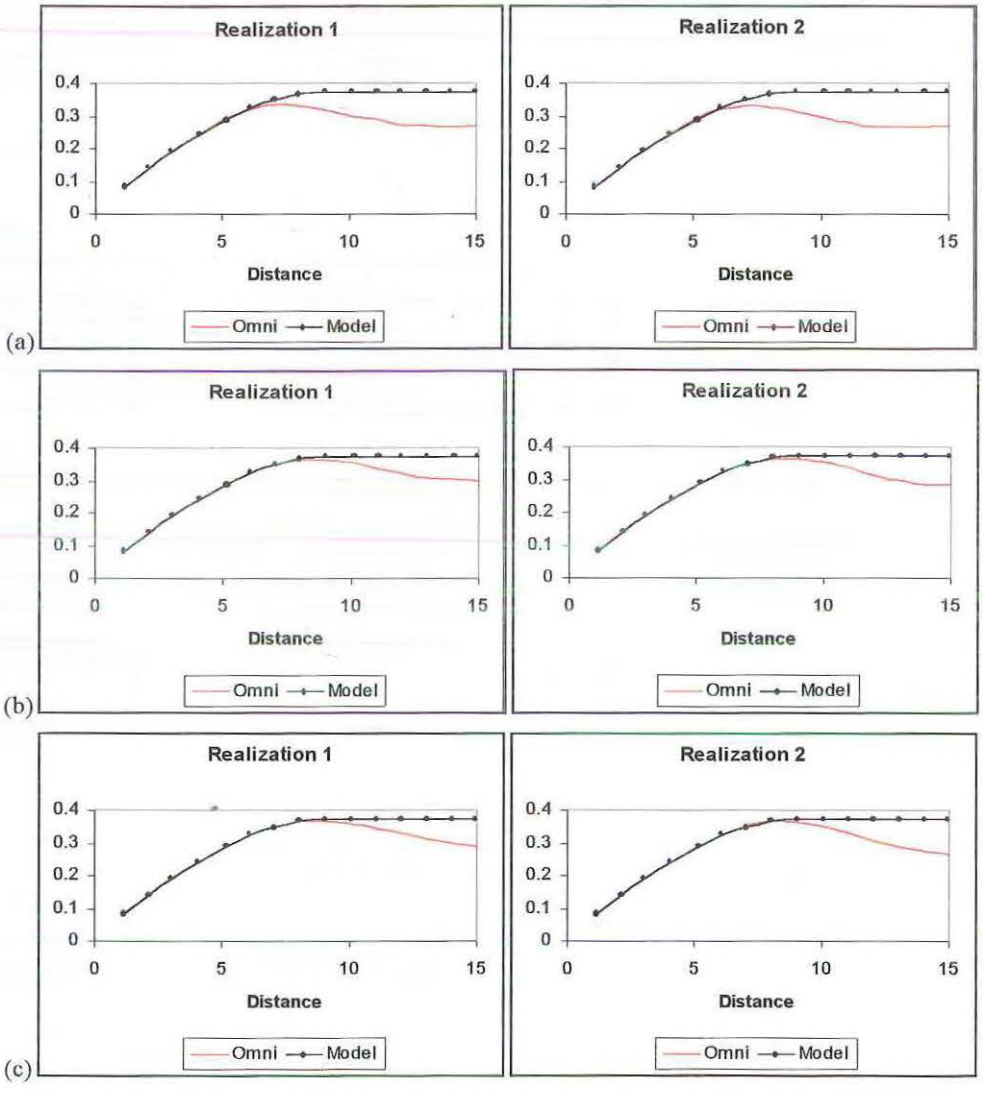

Figure 7.5: The omnidirectional semivariograms for the realizations for nlags equal to 56, 98 and 120.

Using nlags equal to 56 in Figure 7.5a and nlags equal to 98 in Figure 7.5b, the sill of the required model is not reached, but using nlags equal to 120 in Figure 7.5c, the sill is reached. On all occasions, once the maximum is reached, the semivariograms tend to reduce to a lower value. If we were to choose the

'best' results, we would consider using nlags equal to 120 as producing the best results. In this case without including the histogram component in the objective function, we would choose to use the realizations where nlags equals 120.

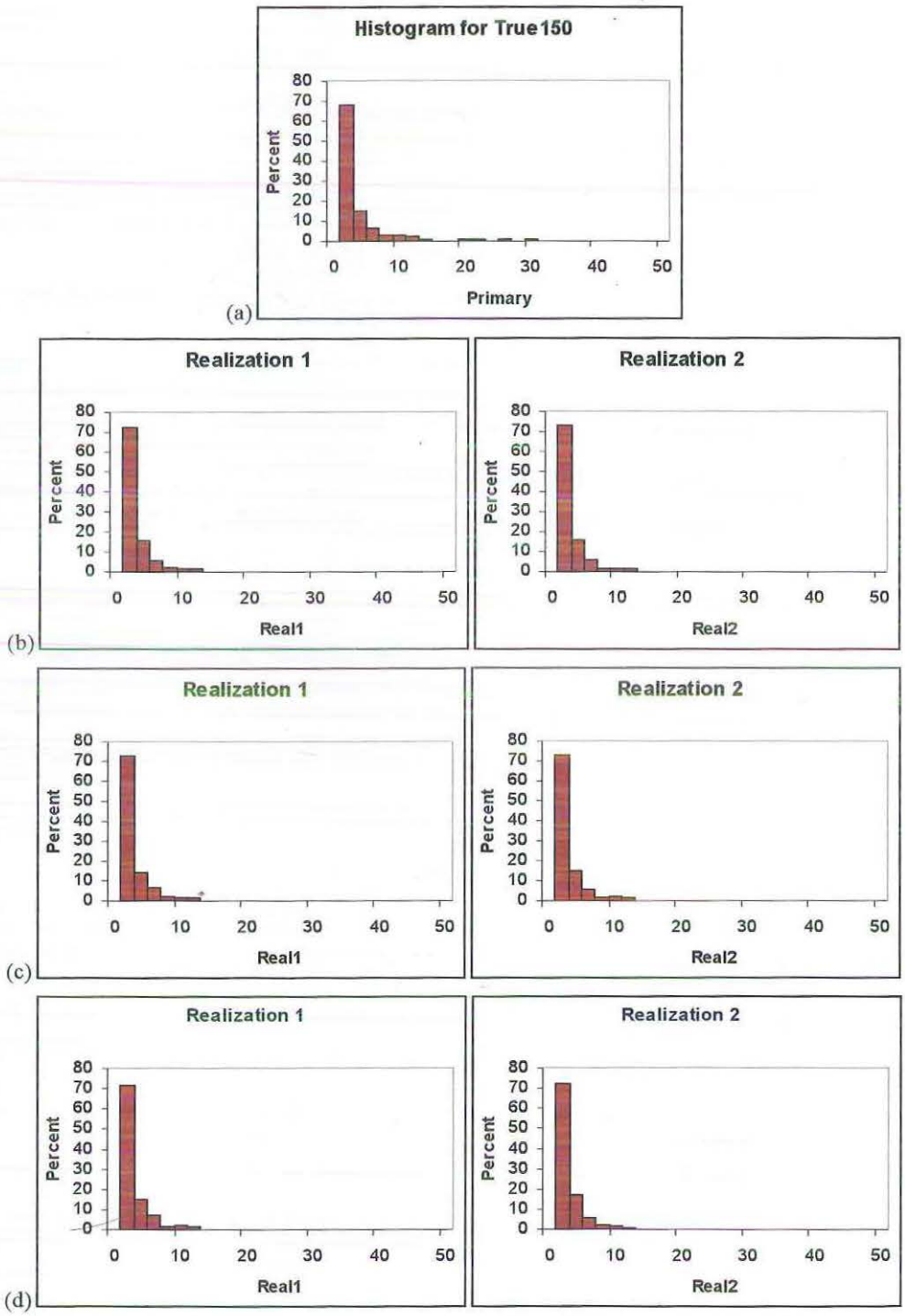

Figure 7.6: The histograms for the realizations for nlags equal to 56 (b), 98 (c) and 120 (d).

In this section, we have not used the histogram component in the objective function. For each of the realization maps in Figures 7.3, we could be interested

in checking how well the histogram for the sample data was used by generating the histograms for the realizations. The histograms for each realization for each nlags are shown in Figures 7.6, together with the histogram for the sample.

The histograms of the realizations in Figures 7.6 are all similar, but when compared with the histogram for the sample in (a), we can see that the histogram for the sample contains more values between the values 20 and 30. While we were not trying to reproduce the sample histogram, it shows that SASIM uses the histogram of the sample when generating the realizations, so we expect an approximate reproduction of the sample histogram.

Berea100.

The SASIM parameter file *BerV.par* uses *Berea100* for Simulated Annealing and can be seen in Appendix C.6. This parameter file sets only the semivariogram component and changes the nlags parameter to 80, 153 and 216. Also, the data have not been log-transformed as we did for *True150*. The maps for the realizations for nlags equal to 80, 153 and 216 are shown in Figures 7.7, respectively, together with the exhaustive and sample data maps.

For the six realizations in Figure 7.7, there is visible banding in the northwest-southeast direction. In the top right-hand corner of the exhaustive data maps in (a), we see a region of high values, but this is not seen in the realizations. This could be caused by the lack of sample observations in that region (see the map of the sample in (a)). In the realizations in Figure 7.7c and Figure 7.7d where nlags equals 153 and 216 respectively, there appears be a small region of high values in the centre. In the regions of low values in the map of the sample, we can see that there are regions of low values in the realizations.

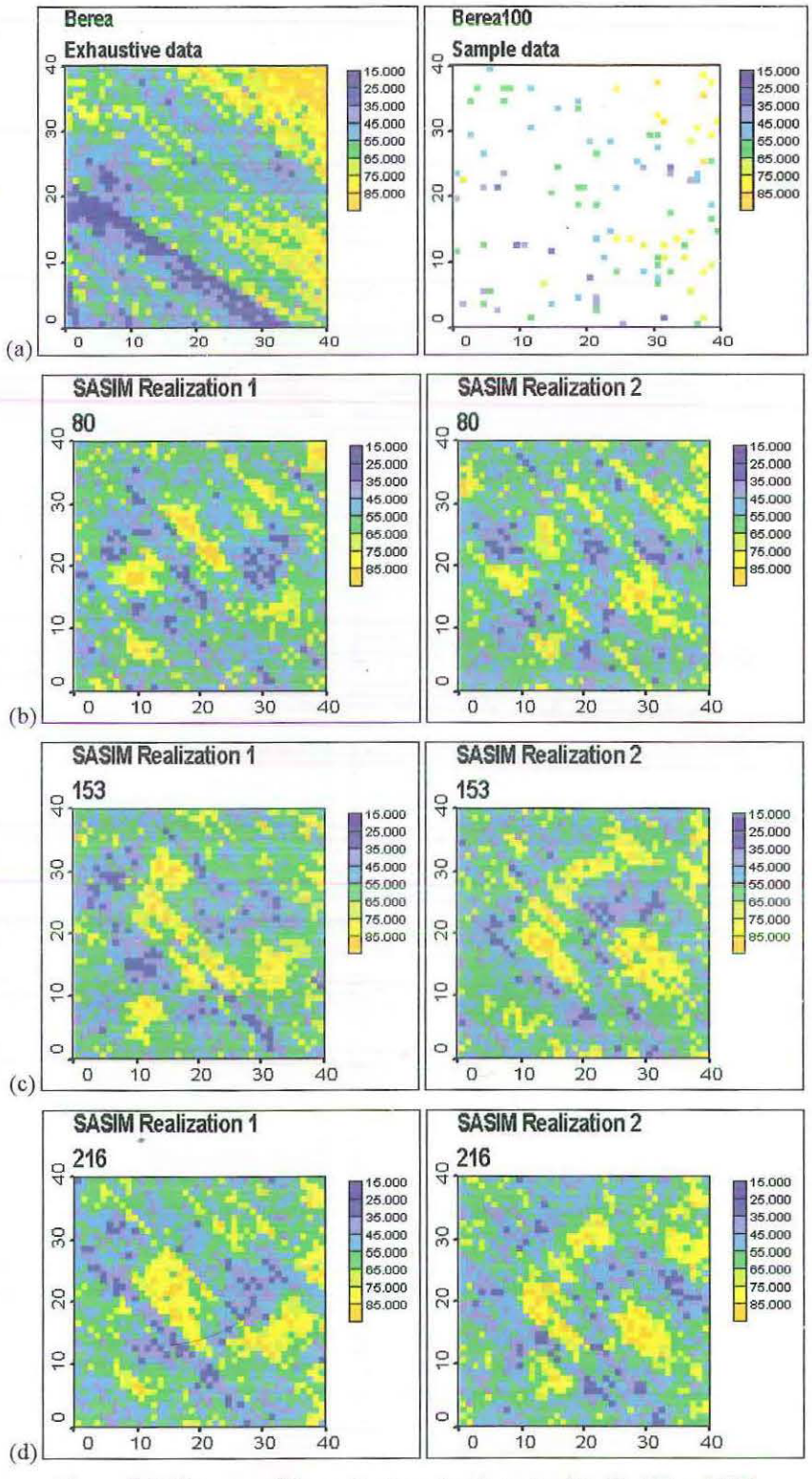

Figure 7.7: The maps of the realizations for nlags equal to 80, 153 and 216.

Using VARMAP, we can generate a semivariogram surface for each realization. The semivariogram surfaces of the realizations for nlags equal to 80, 153 and 216, together with the semivariogram surface generated for *Berea* are in Figures 7.8. In all cases, a lag spacing of 1 with 25 lags is used. Also, we needed

to change the levels from the levels used in Figure 3.6, because the levels for these realizations contain much lower values.

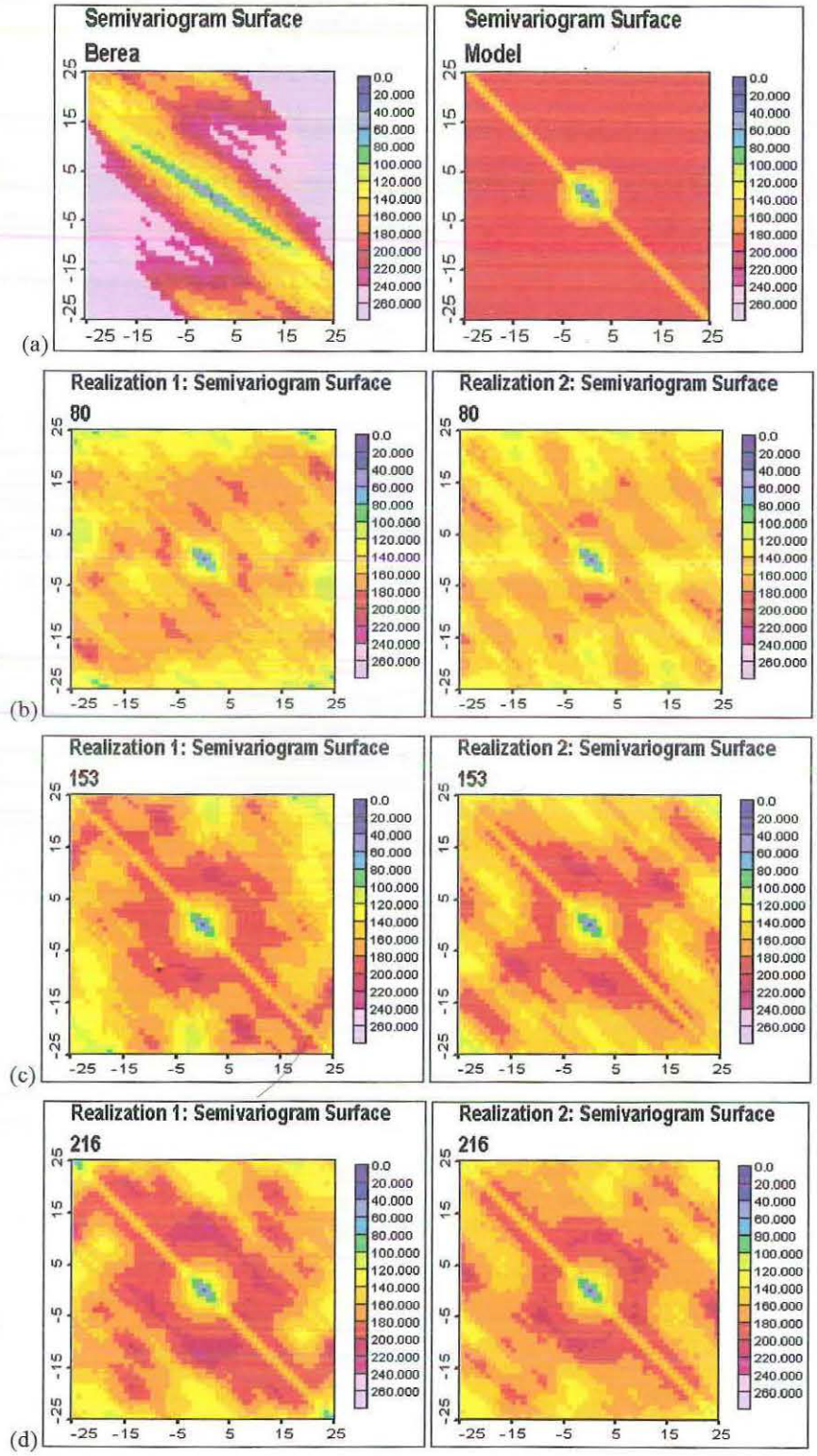

Figure 7.8: Semivariogram surfaces of the realizations for nlags equal to80, 153 and 216.

From Figure 7.8, the anisotropy is approximately in the northwest-southeast direction for the realizations. This direction is different from the semivariogram surface of the exhaustive data in (a) and similar to the semivariogram surface of the sample data.

Since the data are anisotropic, semivariograms in the directions of maximum and minimum continuity are required for each realization. Using GAM, these semivariograms are in Figures 7.9 for the different nlags values, together with the required model.

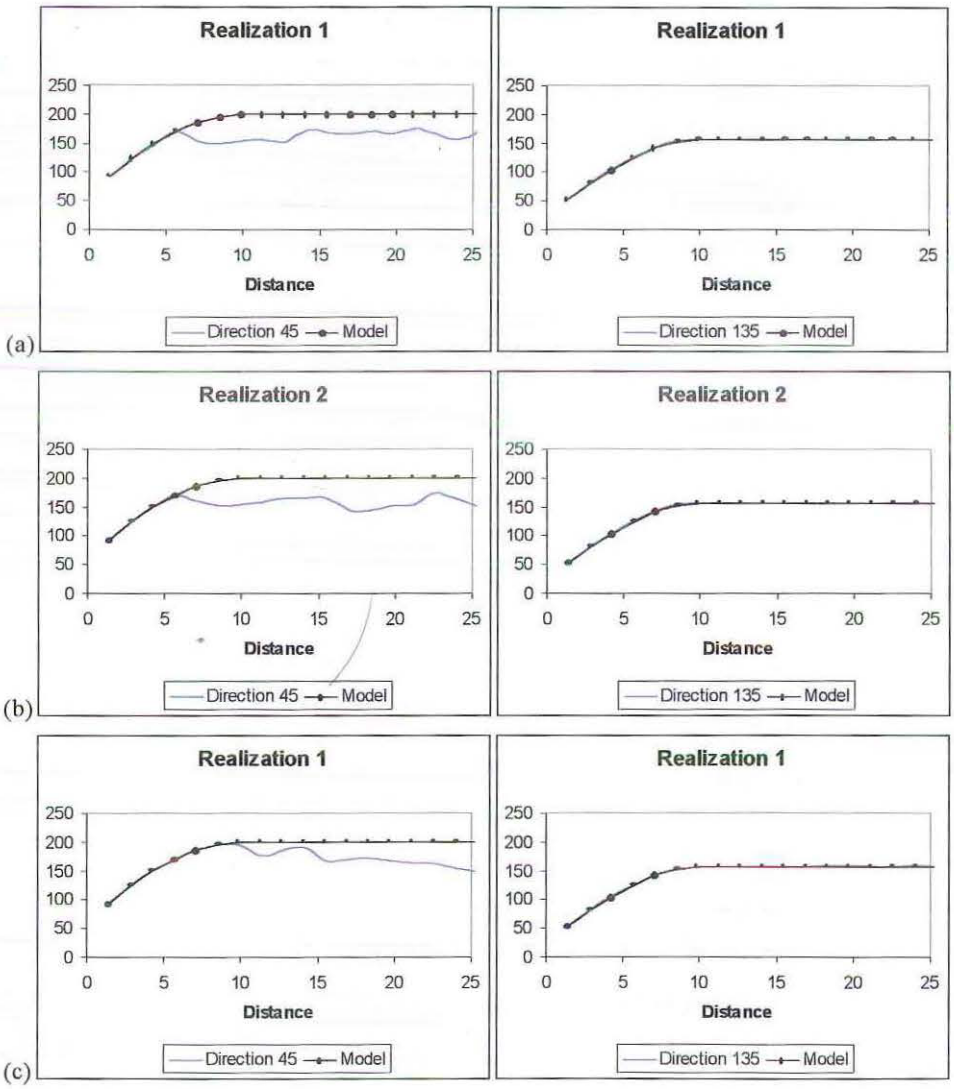

Figure 7.9: The semivariograms of the realization for nlags equal to 80, 153 and 216.

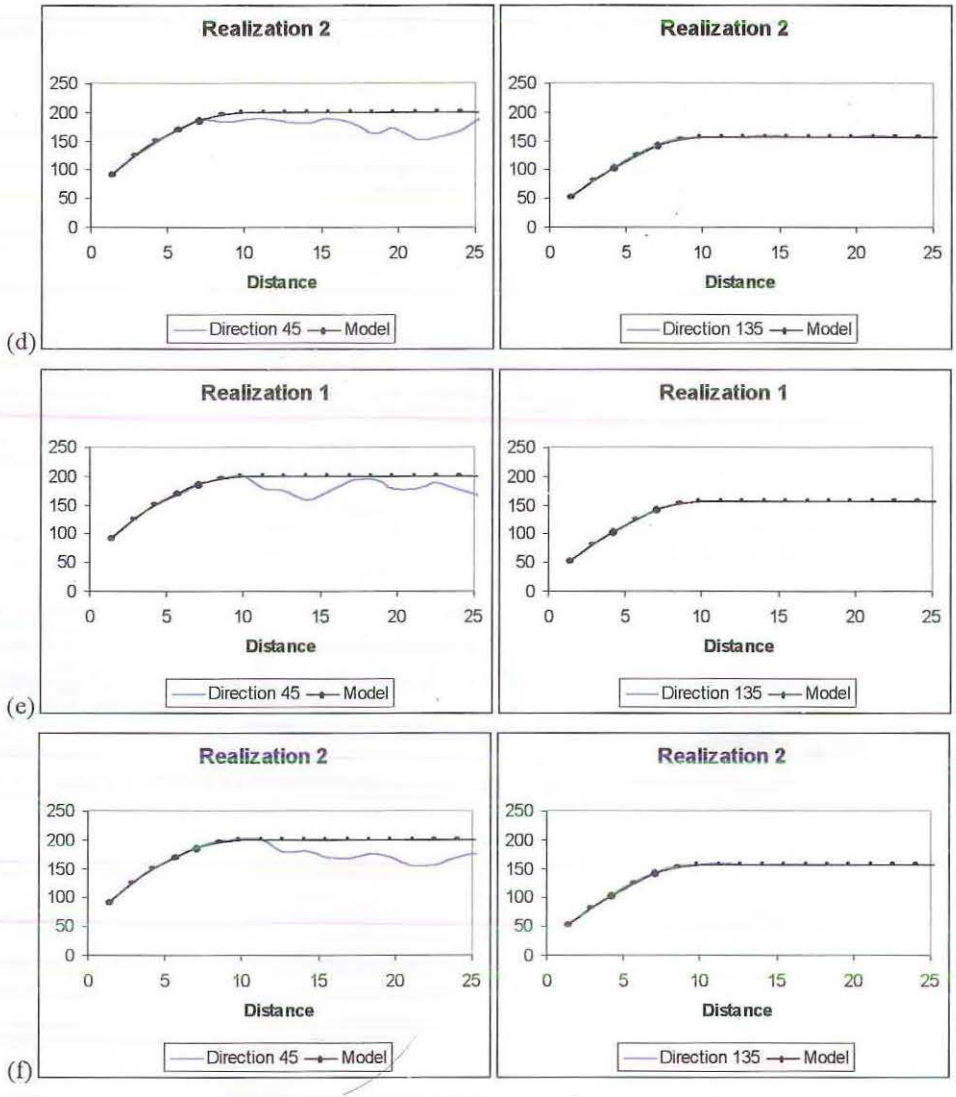

Figure 7.9 continued.

The semivariograms in Figure 7.9 for the realizations with nlags equal to 80 are shown in (a) and (b), for nlags equal to 153 are shown in (c) and (d) and for nlags equal to 216 are shown in (e) and (f). For all realizations in Figure 7.9 the semivariogram model in the direction 135° is almost perfectly reproduced, but the semivariograms in the perpendicular or 45° direction do not reach the sill of the semivariogram model, until nlags is maximum, and once it peaks, it tends to a lower value. The semivariogram in direction 135° at nlags equal to 80 has already reproduced the semivariogram model.

Although, we are not requiring that the sample histogram be reproduced, it is worth noting how similar the histograms of the realizations are to the histogram

of the sample. The histograms of the realizations for nlags equal to 80 (b), 153(c) and 216 (d), together with the histogram for the sample data are in Figures 7.10.

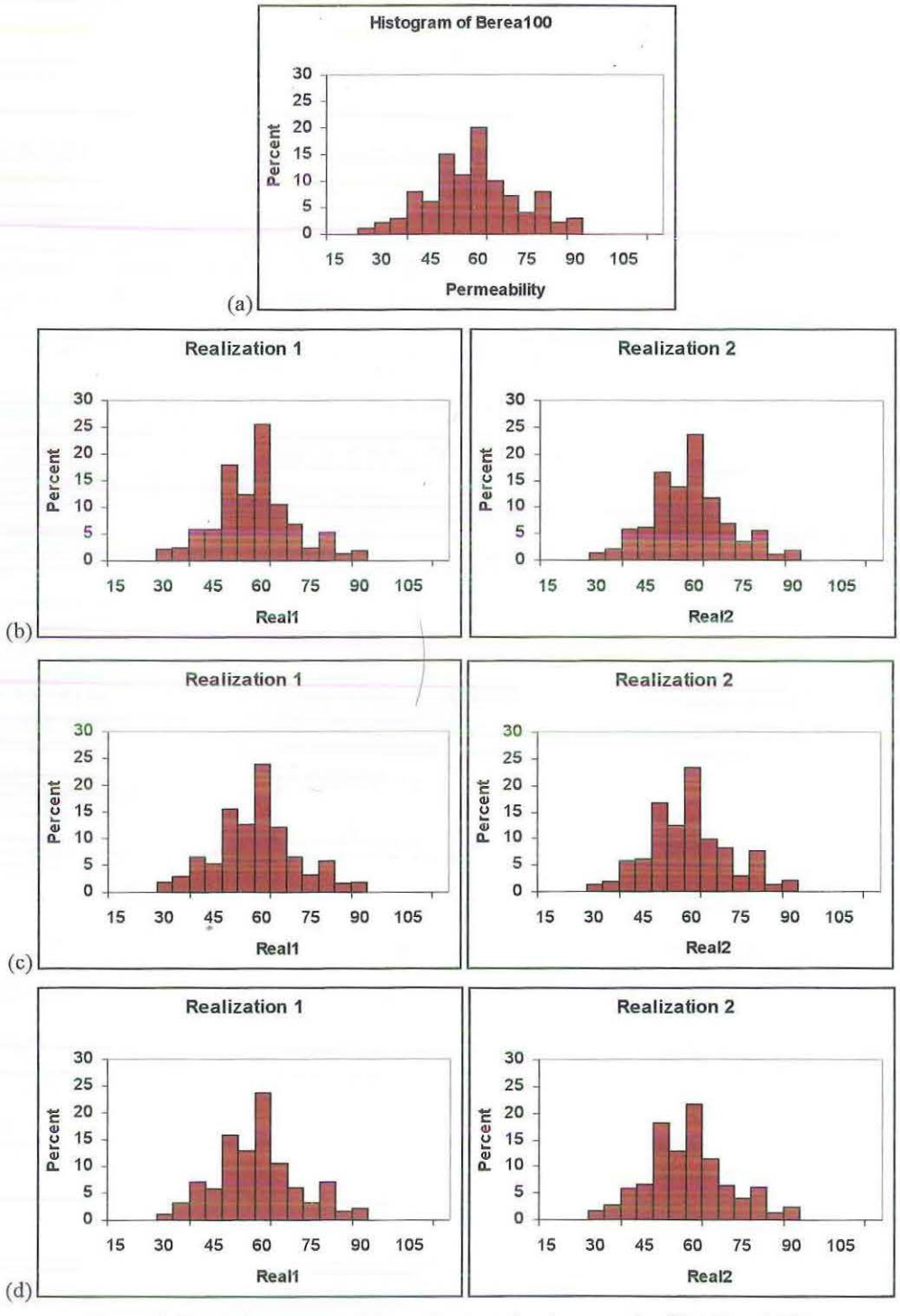

Figure 7.10: The histograms of the realizations for nlags equal to 80, 153 and 216.

Comparing the histogram for the realizations in Figures 7.10 with the histogram of the sample in (a) we can see that they are similar. The number of observations in classes 50 and 60 in the sample has been reproduced in the

histograms of realizations. We can say that the histograms for the realizations have reproduced the histogram of the sample.

7.4 Using both semivariogram and histogram components in SASIM.

True150.

In Section 7.3 we were trying the reproduce only the semivariogram. We saw that the omnidirectional semivariogram for nlags equal to 56 and nlags equal to 98 did not reach the sill, but the sill was reached for nlags equal to 120. We also saw that the reproduction of the histogram was close but not exact. In this section we are trying to improve the realizations in Section 7.3 by adding another prerequisite. That is we want to reproduce not only the semivariogram model, but also the sample histogram. This means we need to add a second component to the objective function in the hope to improve the reproduction of the semivariogram model and histogram. This second component is the histogram. Also, nlags equal to 56, 98 and 120 will be used. Although unlikely, we will keep using nlags equal to 56 just in case the sill of the semivariogram model is reached when we include the histogram.

The second SASIM parameter file *TrVH.par* in Appendix C.6 sets the semivariogram and histogram component parameters; therefore we wish to reproduce the semivariogram model in Section 3.3 and the sample histogram. The weight of the semivariogram component is set to five while the weight of the histogram component is set to one. This parameter file is run using the SASIM executable file and 50 realizations are generated. This results in 50 sets of data, stored in a single column.

The maps for the three realizations for each nlags equal to 56, 98 and 120 are in Figure 7.11, together with the map for *True* and *True150.*

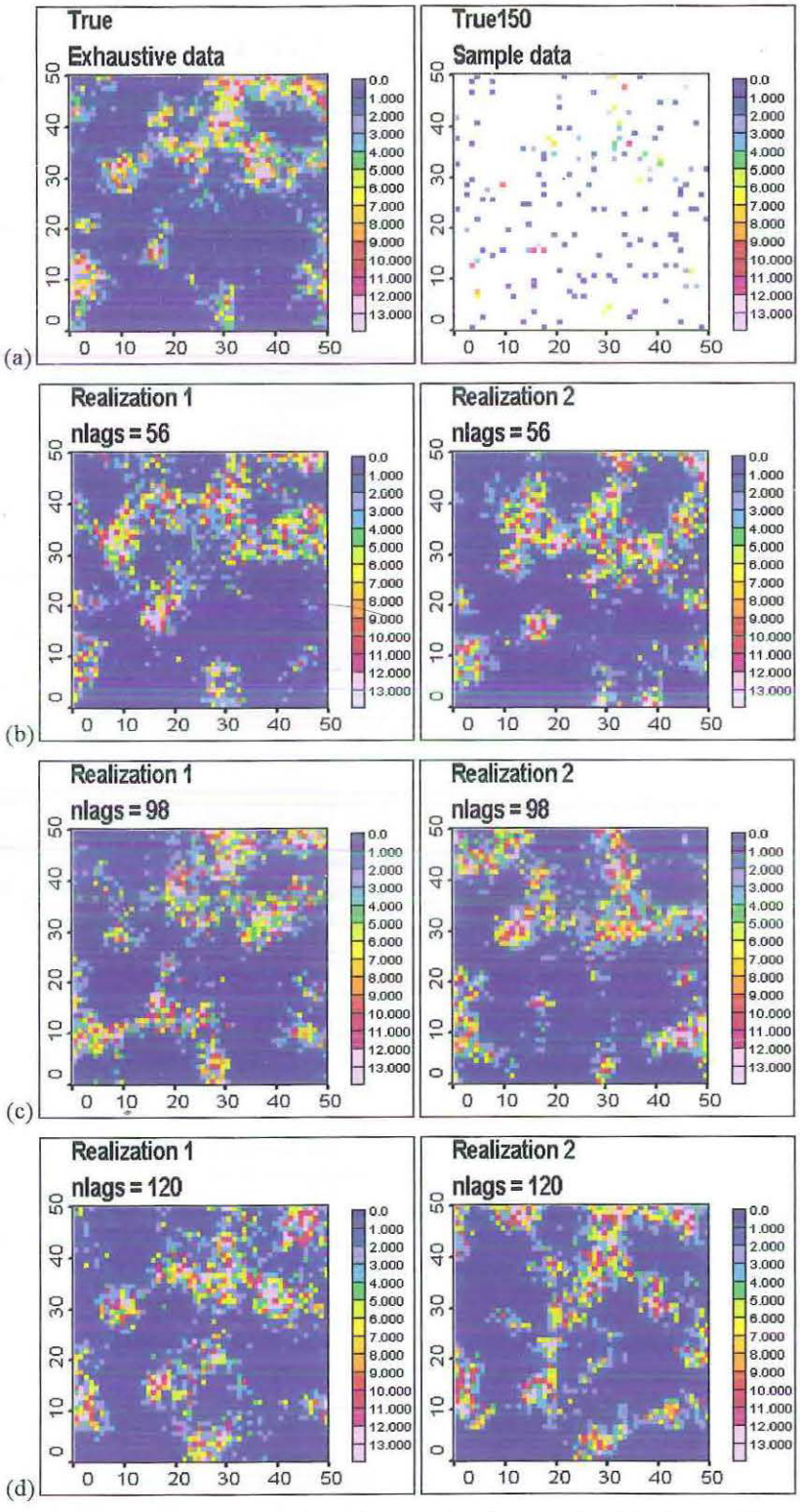

Figure 7.11: The maps of the realizations for nlags equal to 56, 98 and 120.

For each of the realization maps in Figure 7.11, it can be seen that it is similar to the corresponding map for the exhaustive data in (a), but there appears to be one region at approximately (30, 20) in the realizations for nlags equal to

120, that consists of high values not seen in the exhaustive data. The regions of high values seen in the map for the sample data in (b) have been reflected in the realizations. At approximately (30, 20) the map for the sample has no sample points and so in the some of the maps for the realizations this region of values has been created. It seems that the realizations for nlags equal to 98 are the best representations for the exhaustive data.

Remember that we are attempting to reproduce both histograms and semivariograms; the histograms for each realization for each nlags are shown in Figures 7.12, together with the histogram of the sample. The histograms for nlags equal to 56 are shown in (b), for nlags equal to 98 in (c) and for nlags equal to 120 in (d).

The histograms in Figure 7.12 are very similar to each other, as well as to that of the sample data in (a). It appears that the histogram for the sample was reproduced successfully in the realizations. Comparing these histograms with those in Figure 7.6 in Section 7.3, we can see that the numbers of values between 20 and 30 have increased.

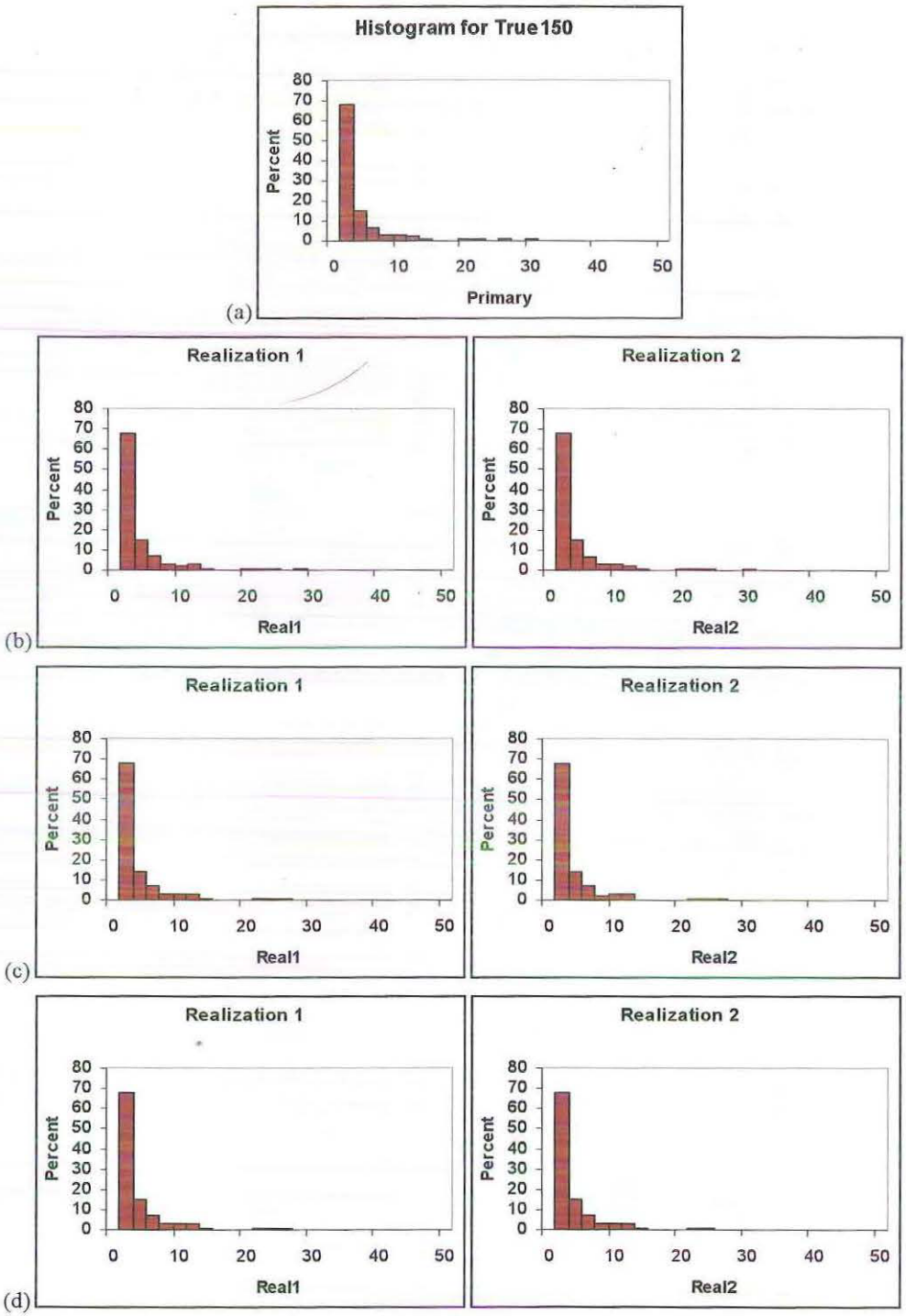

Figure 7.12: The histograms for the realizations for nlags equal to 56, 98 and 120.

Using the VARMAP parameter file we can generate semivariogram surfaces of the logarithmic data, and map them in 3Plot. For all semivariogram surfaces in Figure 7.13, a lag spacing of 1 with number of lags equal to 30, is used.

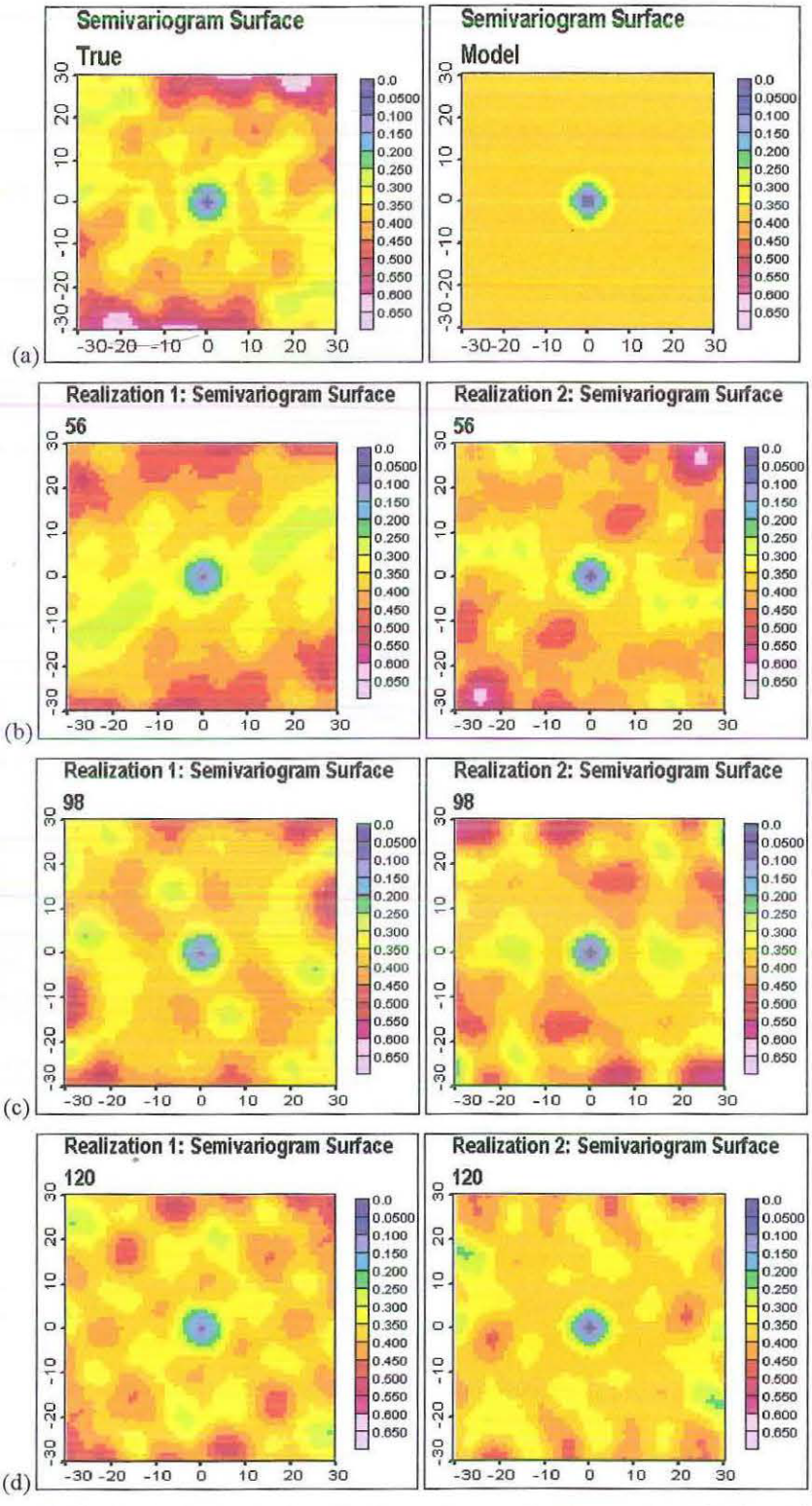

Figure 7.13: Semivariogram surfaces for the realizations for nlags equal to 56, 98 and 120.

The semivariogram surfaces in Figure 7.13 all show that the isotropy has been reproduced in the realization as we saw in the semivariogram surface for the exhaustive in (a) and sample data in (b). Furthermore, the regions of high values around the edge of the grid in (a) are also seen in the realizations.

For each realization, the omnidirectional semivariograms are generated using GAMV and compared with the semivariogram model we wish to reproduce. These can be seen in Figures 7.14.

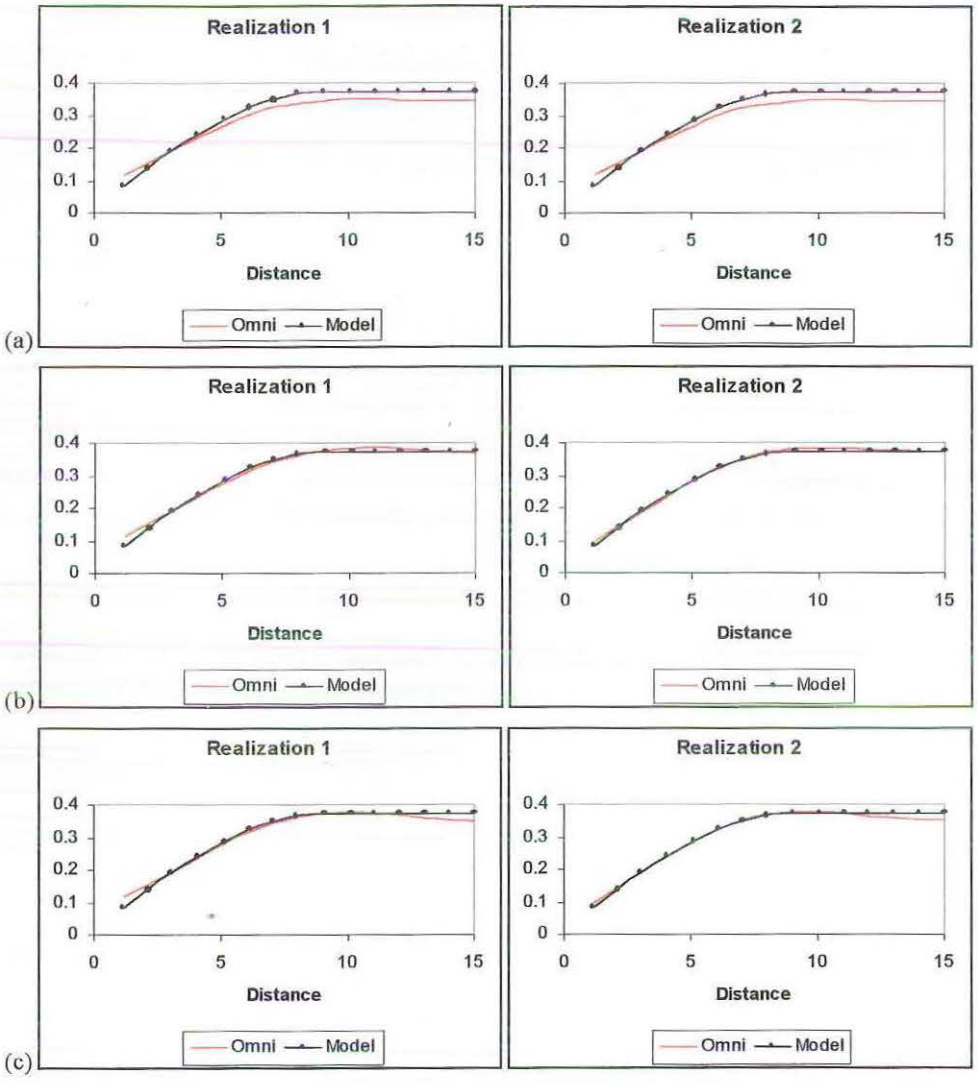

Figure 7.14: The omnidirectional semivariograms for the realizations for nlags equal to 56, 98 and 120 and the required model.

When using nlags equal to 56, seen in Figure 7.14a, the sill of the model is not reached. But the sill is reached when using nlags equal to 98, seen in Figure 7.14b, or nlags equal to 120 in Figure 7.14c. In most instances, the realizations tend to overestimate the sill, especially when nlags equal to 56. For nlags equal to 98 and 120, there is no substantial difference. So when including the factor of processing time, we will use nlags equal to 98.

Berea100.

The second SASIM parameter file *BerVH.par* in Appendix C.6, sets the histogram and semivariogram components. It also changes the nlags parameter to 80, 153 and 216. This parameter file attempts to reproduce the semivariogram model in Section 3.4 and the sample histogram. The maps of the realizations are shown in Figures 7.15, together with the maps of the exhaustive and the sample data.

The maps of the realizations in Figures 7.15 all show banding in the northwest-southeast direction as seen in the map of the exhaustive data in (a). Also, the region of high values in the top right-hand corner of the exhaustive data map is also seen in the realizations. Although the number of high values in the top right-hand corner in the map of the sample in (b) is limited, there is a region of high values in the maps of the realizations. The regions of low values in the map of the sample are also seen in the map of the realizations.

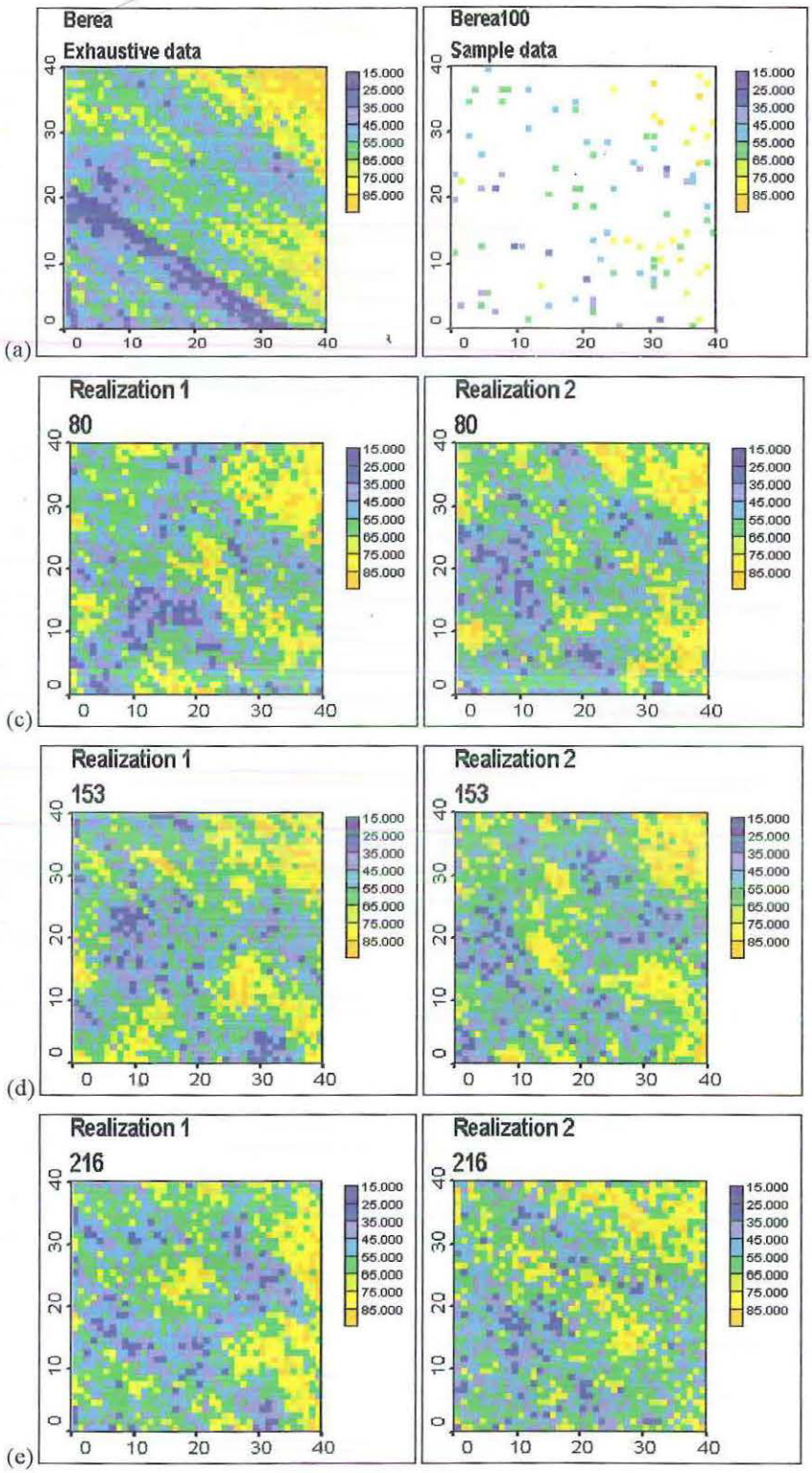

Figure 7.15: The map of the realizations for nlags equal to 80, 153 and 216.

The histograms of the realizations, together with the histogram of the sample, are seen in Figures 7.16. The histograms for the realizations generated for nlags equal to 80, 153, 216 are shown in (b), (c) and (d) respectively.

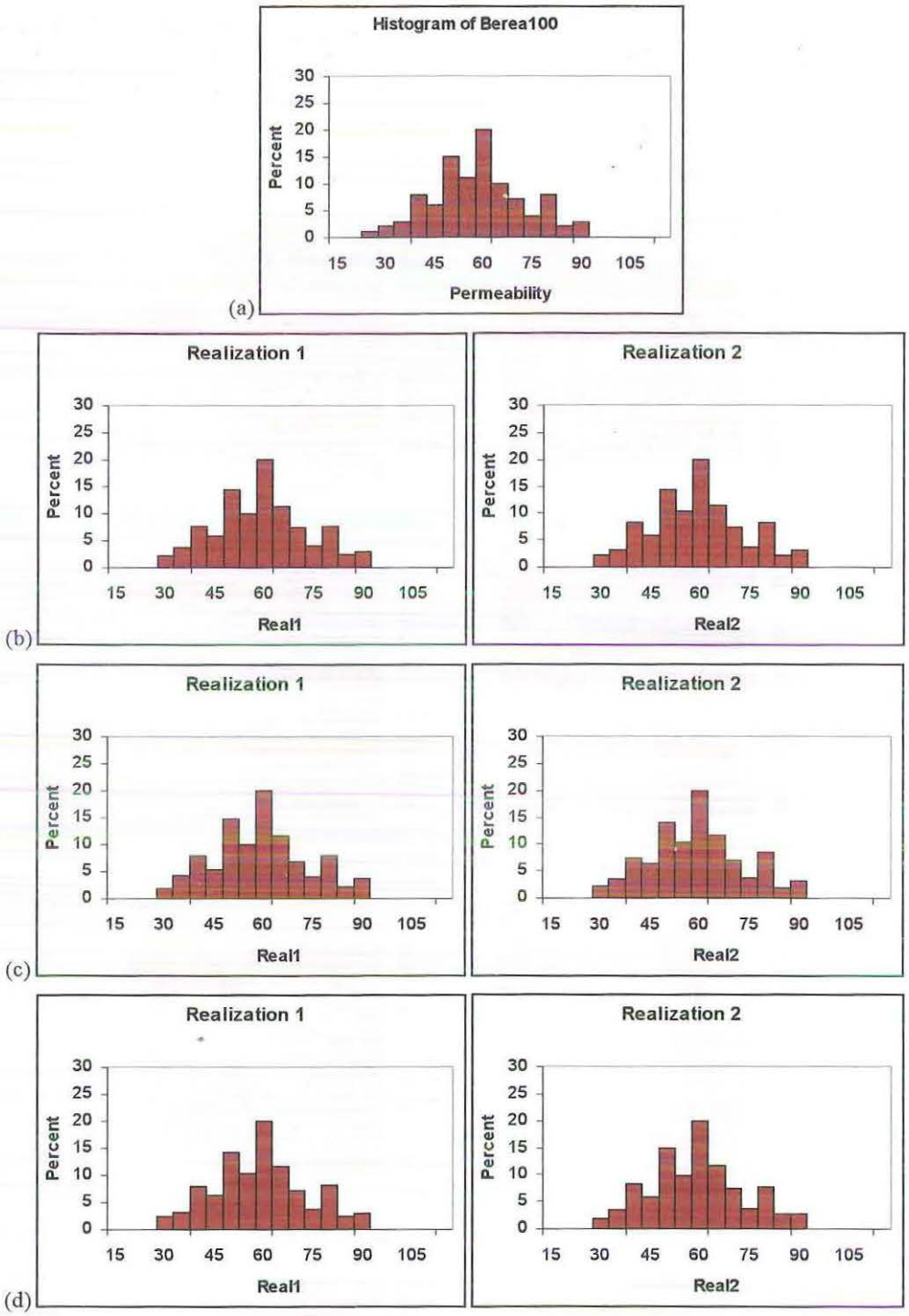

Figure 7.16: The histograms of the realizations for nlags equal to 80, 153 and 216.

The histograms for the realizations in Figures 7.16 are all similar when compared to the histogram for the sample in (a). All realizations are normally distributed, with a high number of values around 60. These are reproduced from the sample histogram.

The semivariogram surfaces generated in VARMAP for the realizations for different nlags values are in Figures 7.17, together with the semivariogram surface for *Berea*.

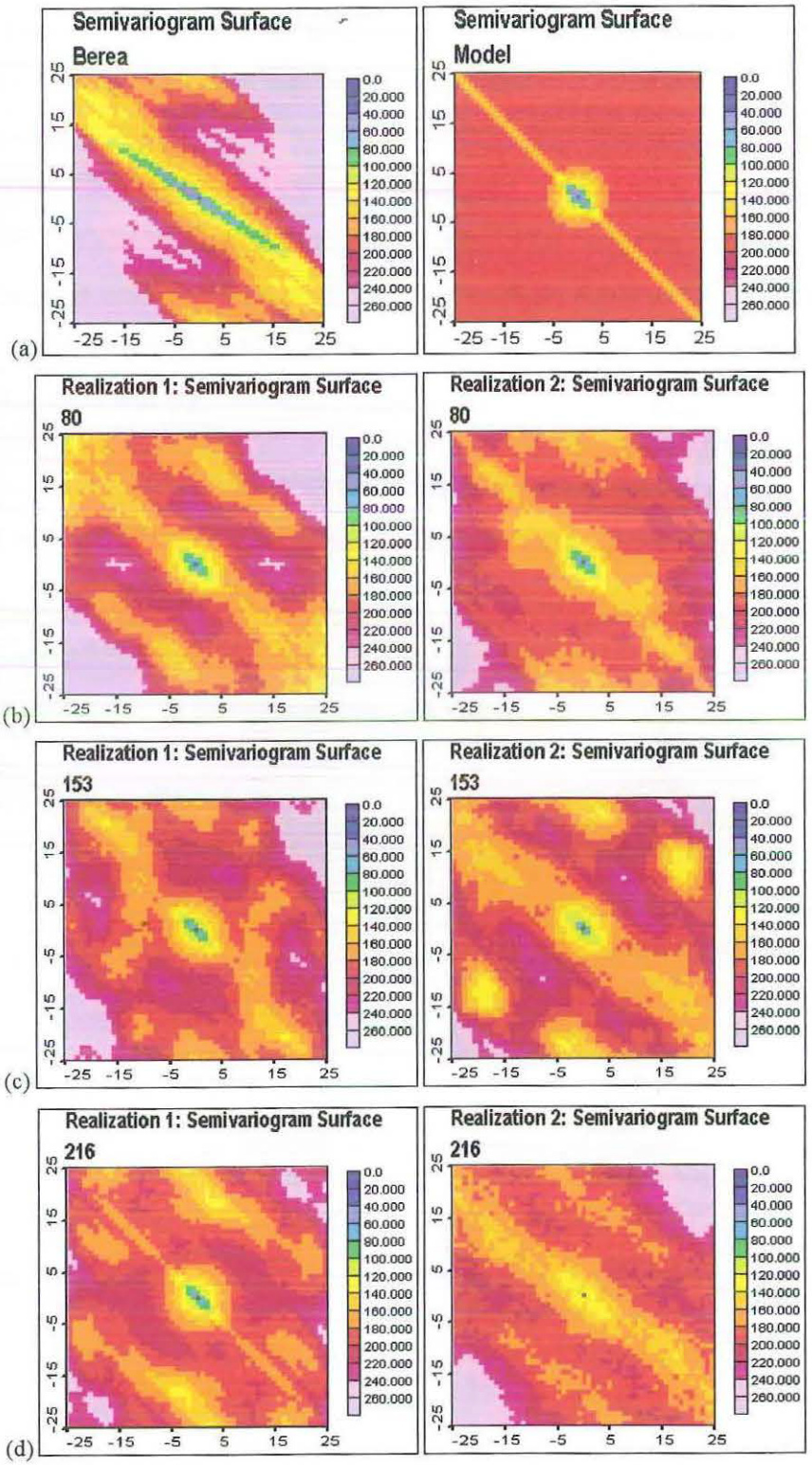

Figure 7.17: Semivariogram surfaces of the realizations for nlags equal to 80, 153 and 216.

The semivariogram surfaces for the realizations in Figures 7.17, all show anisotropy in the northwest-southeast direction. This direction is visibly different in the semivariogram surface for *Berea* in (a), but we use the semivariogram surface for the sample. Each semivariogram surface for the realizations has greater spatial variability than the previous semivariogram surfaces in Figure 7.8 where only a semivariogram was being reproduced.

The semivariograms for the realizations for the different nlags values generated in GAM are shown in Figure 7.18 together with the target semivariogram model. The semivariograms for nlags equal to 80 are shown in (a) and (b), for nlags equal to 153 are shown in (c) and (d) and for nlags equal to 216 are shown in (e) and (f).

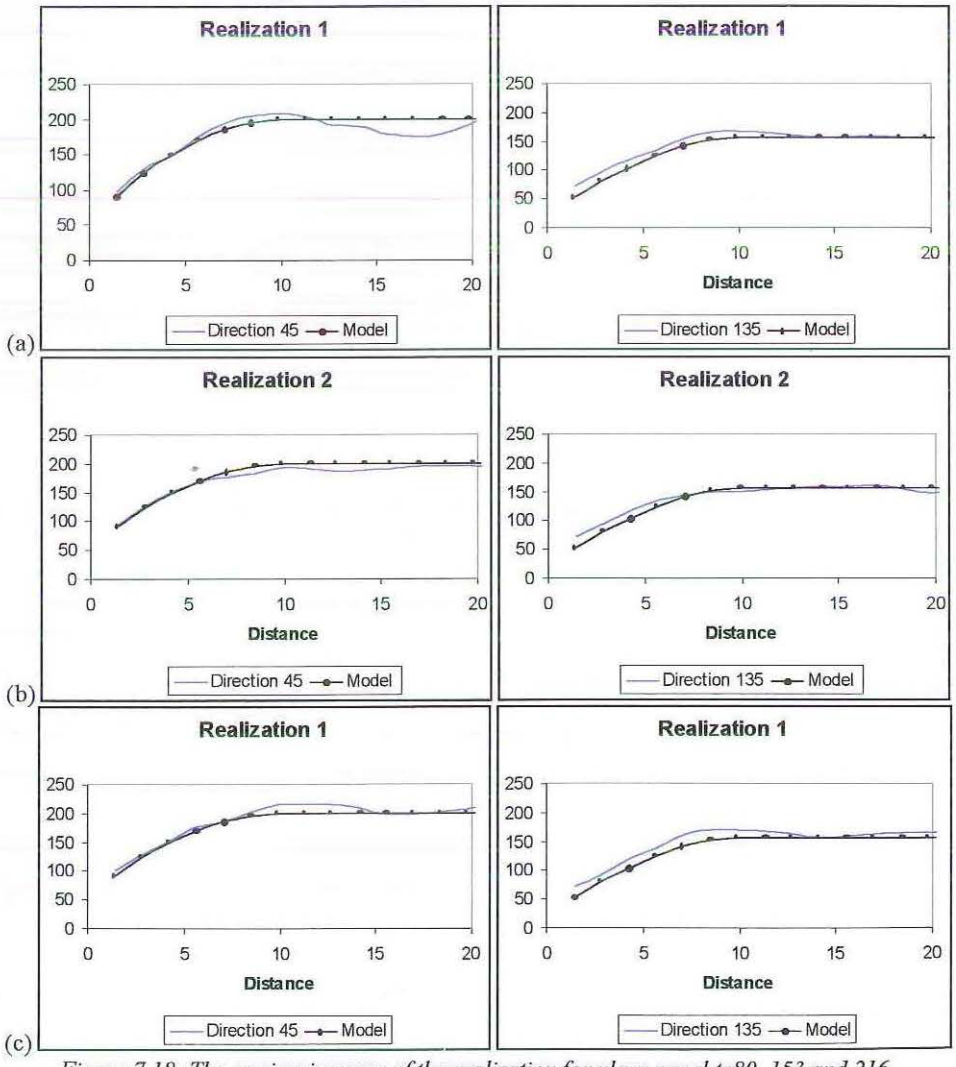

Figure 7.18: The semivariograms of the realization for nlags equal to80, 153 and 216.

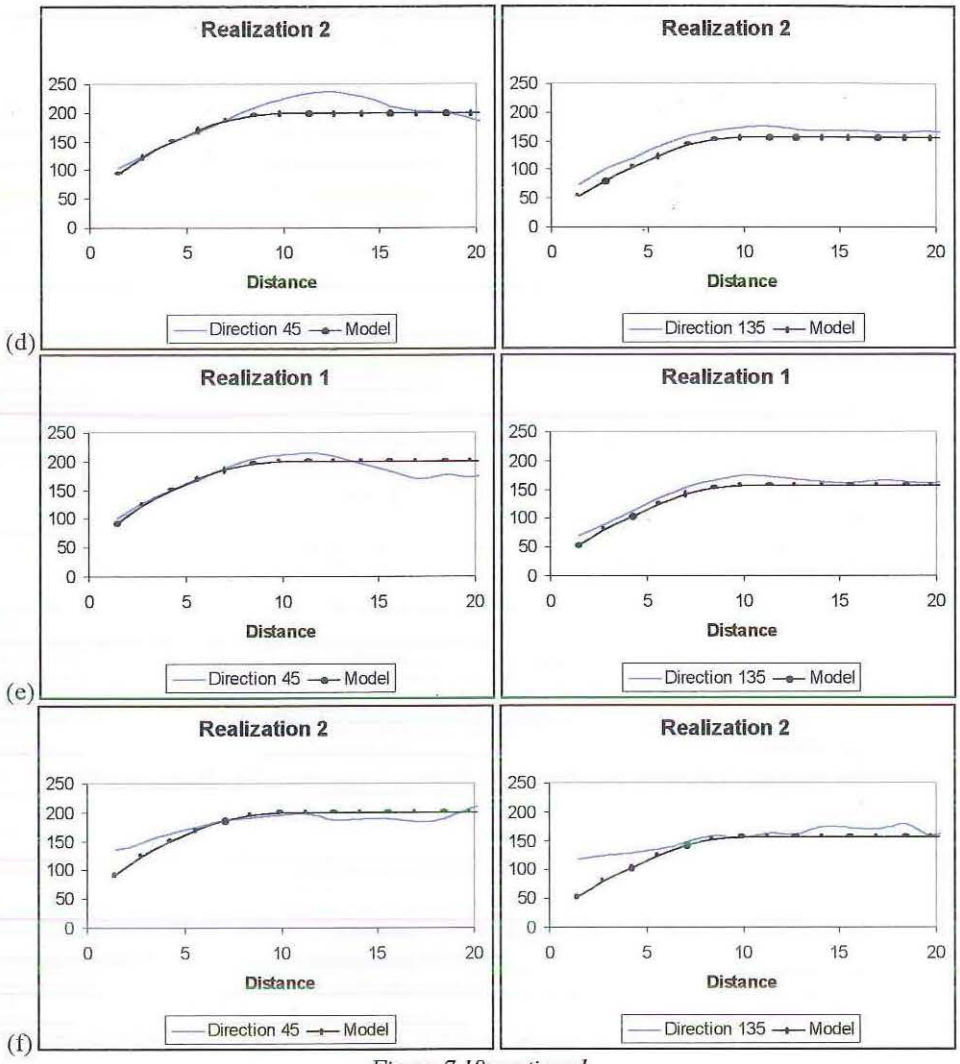

Figure 7.18 continued.

Some of the semivariograms for the realizations in the direction 135° are similar to the required model, but the semivariograms in direction 135° were much better reproduced in Figure 7.9 where only the semivariogram component was set in the objective function. The semivariograms in the direction 45° are not as similar, but are generally better reproduced than the semivariograms in direction 45° in Figure 7.9. Overall, using the different nlags, the semivariograms in Figures 7.18 have been reproduced reasonably well. Although using nlags equal to 80 reproduces similar semivariograms, it appears to be too small because the semivariograms in direction 45°, the values are either too high or too low. Using the maximum number of nlags (216) sometimes generates worse realizations than either nlags equal to 80 or 153. It seems that using nlags equal to 153 produces better realizations than if nlags equals 80 or 216.

7.5 Changing the weight of the semivariogram component.

True150.

In Section 7.4 when the SASIM parameter file *TrVH.par* which included the components for the semivariogram and histogram, it was seen that the best results were obtained when we had nlags equal to 98 and 120, where there was no major difference. Therefore we continued to use nlags equal to 98 because the higher nlags is, the more terms need to be summed and the more time is taken to generate the realizations. For this section we continue to use the SASIM parameter file *TrVH.par* in Appendix C.6, but test the effect if we change the weight (line two) of the semivariogram component. Earlier we used a weight of five for the semivariogram component, but we will now explore the effect of weights one and three for this component.

The first pair of realizations uses the weight one for the semivariogram component. This means that the semivariogram and histogram have equal contribution to the realizations. The next pair uses the weight of the semivariogram component as three. The realization maps for the weight equal to one and three are shown in Figure 7.19, together with the maps for the exhaustive and sample data sets.

These maps of the realizations with weights for the semivariogram component equal to one and three in Figure 7.19 consist of more high values than the map of the exhaustive data shown in (a). It appears that by looking at the maps of the realizations, the semivariograms need more influence than the histogram in reproducing the realizations. The region of high values seen in the map for the sample data in (b) can be seen in the maps of the realizations.

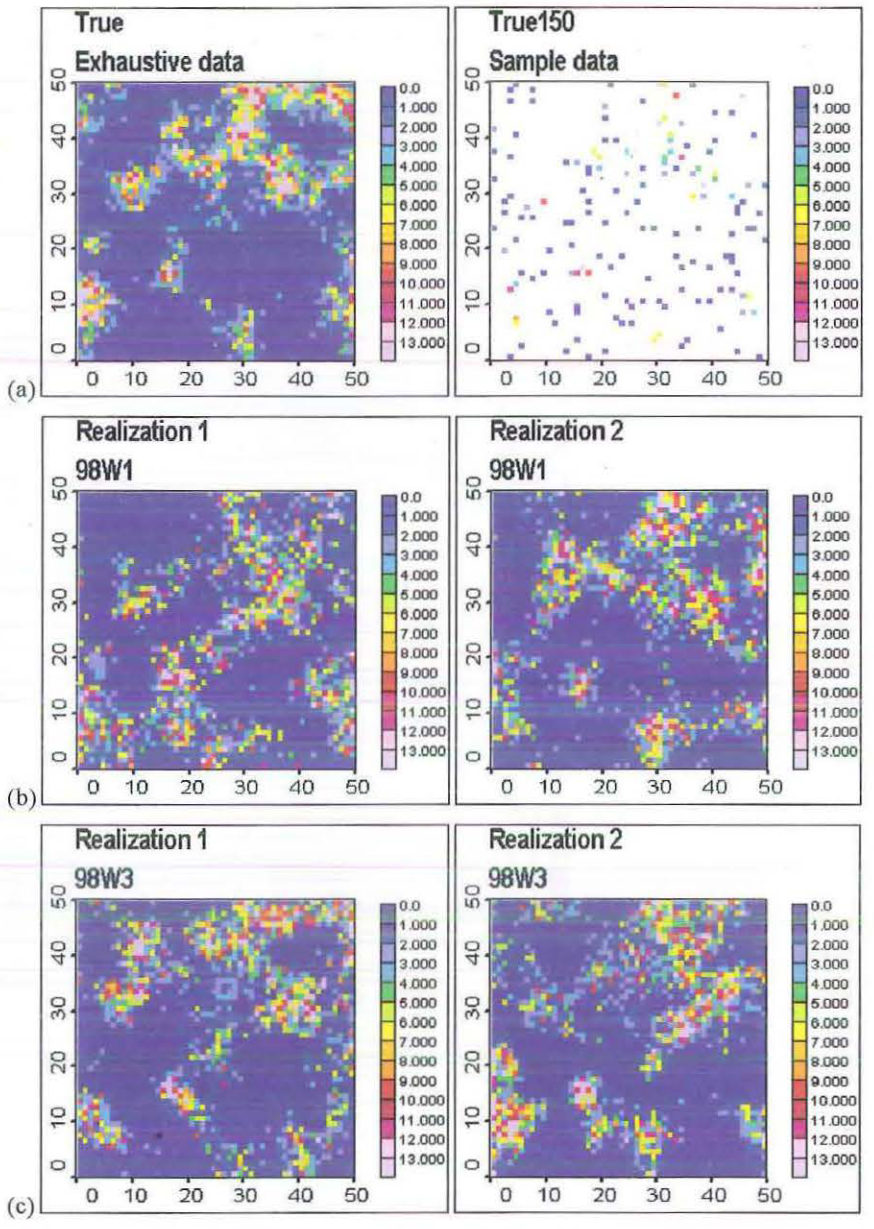

Figure 7.19: Realization maps for nlags equal to 98 and semivariogram weight equal to one and three.

To check if the histograms were reproduced, Figure 7.20 shows the histograms of the realizations, together with the histogram of the sample data. The histograms for the realizations with the weight of the semivariogram component equal to one are shown in (b) and for the weight of the semivariogram component equal to three are shown in (c).

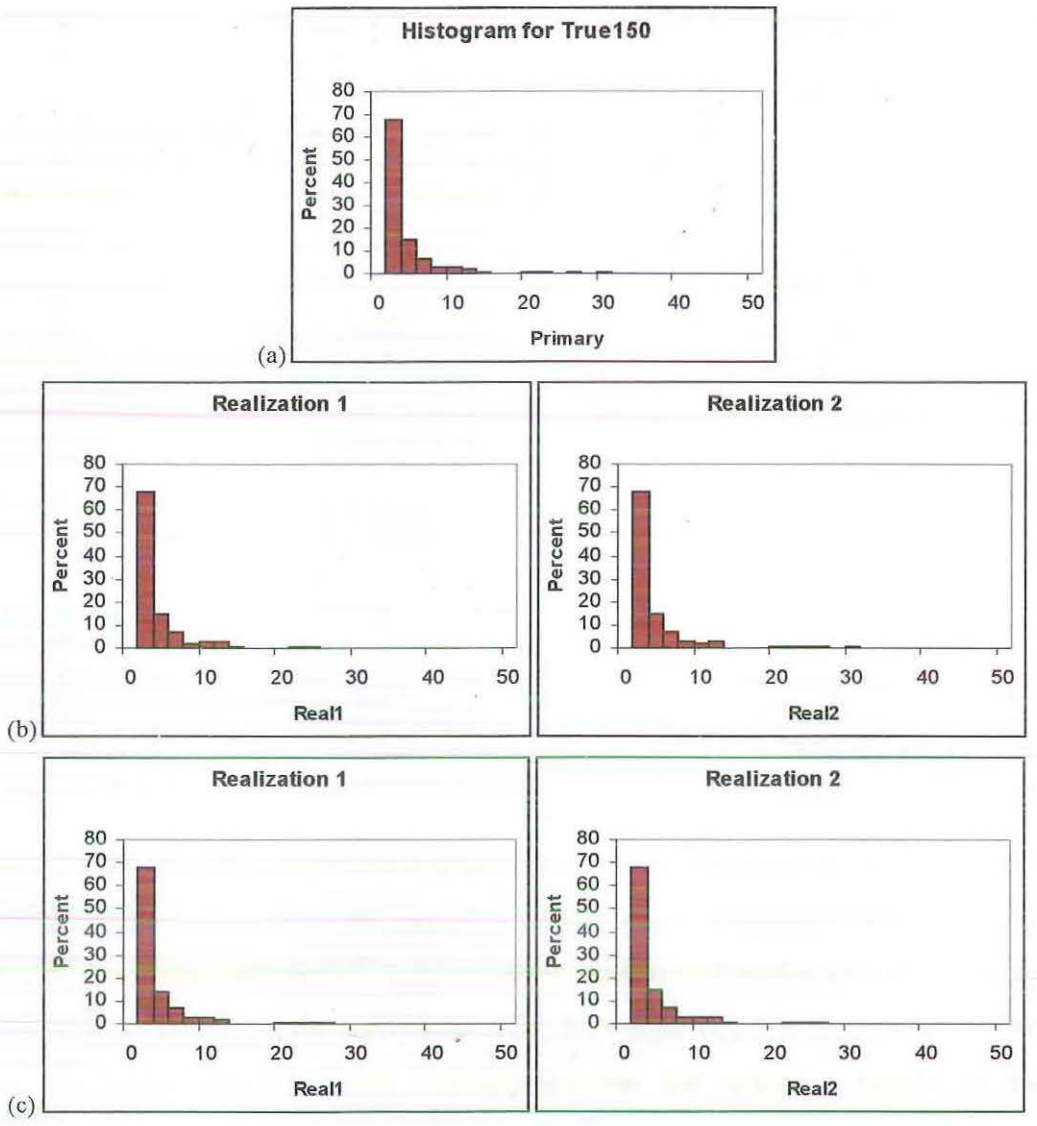

Figure 7.20: The histograms for the realizations.

The histograms for the realizations in Figure 7.20 clearly match the histogram of the sample data in (a), and thus the histograms have been reproduced successfully.

The semivariograms for the two realizations for each weight, together with the required model are shown in Figure 7.21. The semivariograms for the realizations generated with a weight equal to one are shown in (a) and for weight equal to three are shown in (b).

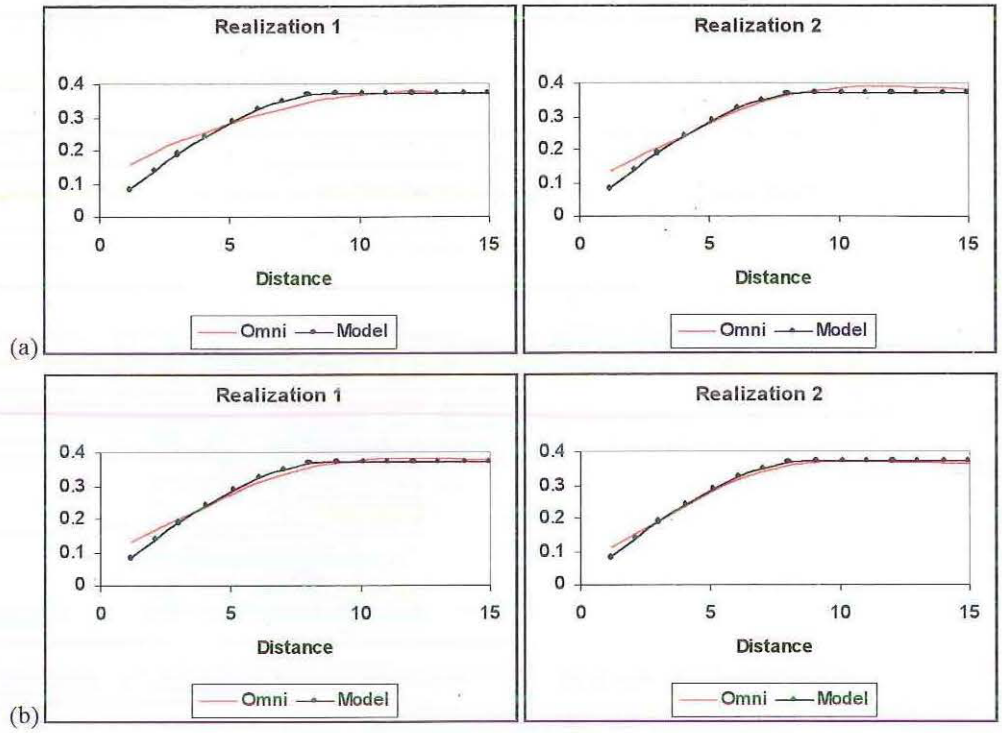

Figure 7.21: The omnidirectional semivariograms for the realizations with the required model.

The omnidirectional semivariograms for the realizations in Figure 7.21 have a higher nugget value and appear to have a longer range. The semivariogram model has been reproduced approximately, but the reproduction is better for weight five. However, the reproduction for the weight equal to three is better than for the weight equal to one. It appears that the optimum weight of the semivariogram component is five.

Berea100.

In Section 7.4 when we used the semivariogram and histogram components in the SASIM parameter file *BerVH.par*, we found that by using the maximum number for nlags of 216 we were able to get the best reproduction of both the histogram and semivariograms in the directions 45° and 135°. Furthermore, we also set the weight of the semivariogram component to the maximum of five. In this section, as we did for *True150* previously, we are changing the weight from five to one and three, to see what impact this has on the realizations.

The first pair of realizations use the weight as one, giving equal contribution to the objective function for the histogram and the semivariogram, and the second pair is for the weight of three. The maps for these realizations are in Figure 7.22, together with the maps for the exhaustive and sample data.

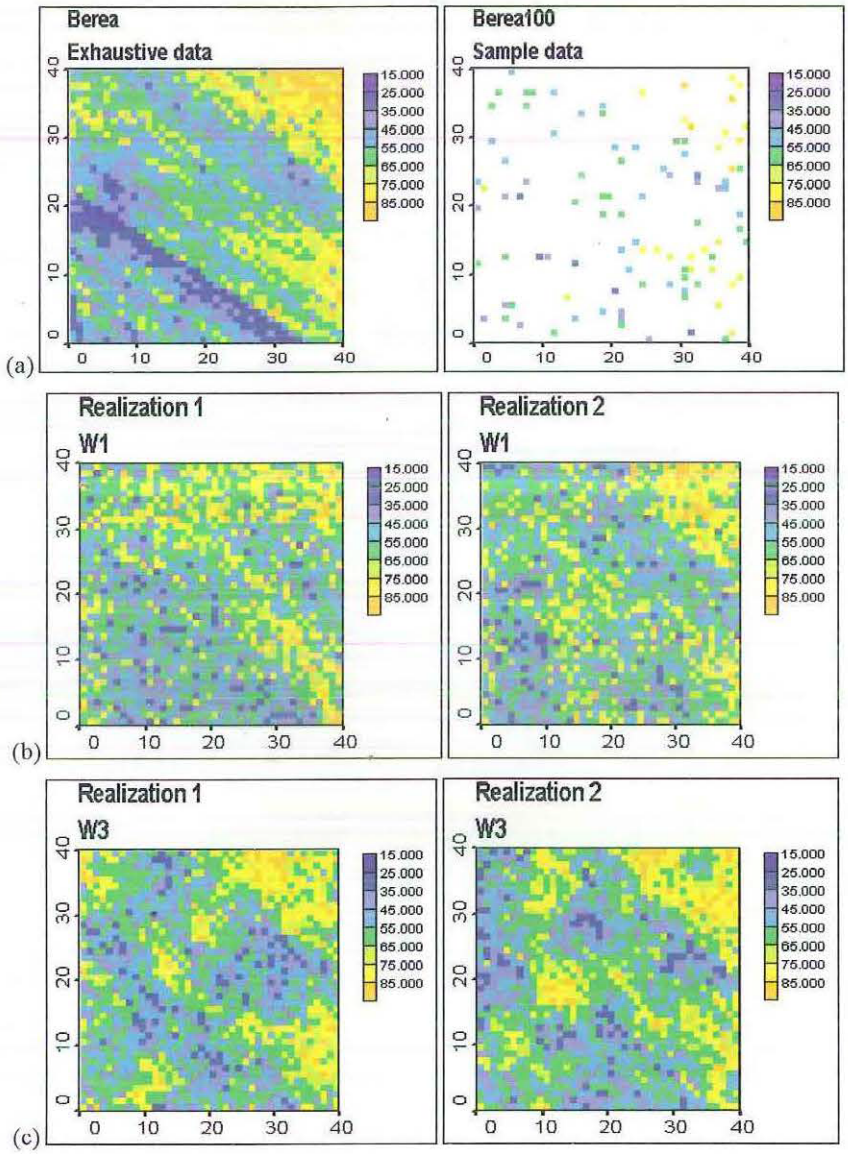

Figure 7.22: The realization maps for nlags equal to 216 and semivariogram weight equal to one and three.

The maps of the realizations in Figure 7.22 when compared to the exhaustive data map in (a), show that there is a region in the top right hand corner of high values, but in the realizations there are also several low values in this region. Also, the realizations for weight equal to one do not appear to have banding as seen in the map of *Berea*. But there appears to be banding in the northwest-southeast direction in the realizations with a weight of three, which is

seen in (a). These realizations also consist of regions of high and low values that are seen in the map of the sample in (a).

The histograms for the realizations are in Figure 7.23, together with the histogram for the sample data. The histograms in (b) are for the semivariogram component equal to one and the histograms in (c) are for the weight equal to three.

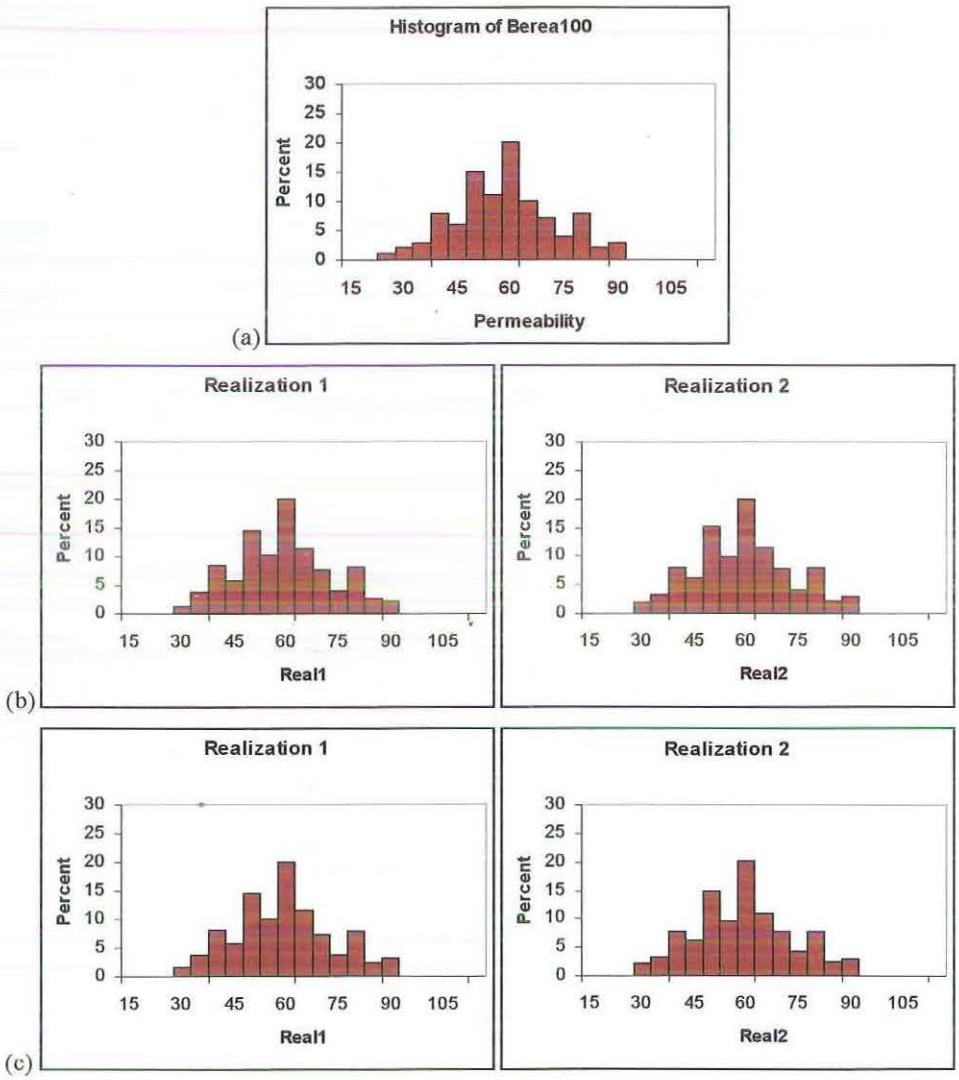

Figure 7.23: The histograms for the realizations.

The histograms of the realizations in Figure 7.23 are similar to the histogram of the sample in (a). The realization data appear to be normally distributed with high numbers of observations at classes 50 and 60. Therefore the histograms of the realization are reproduced from the histogram of the sample.

The semivariograms for the realizations, with the required semivariogram model in black, are shown in Figure 7.24. The semivariograms for the semivariogram component set to one are shown in (a) and (b) and for the weight set to three are shown in (c) and (d).

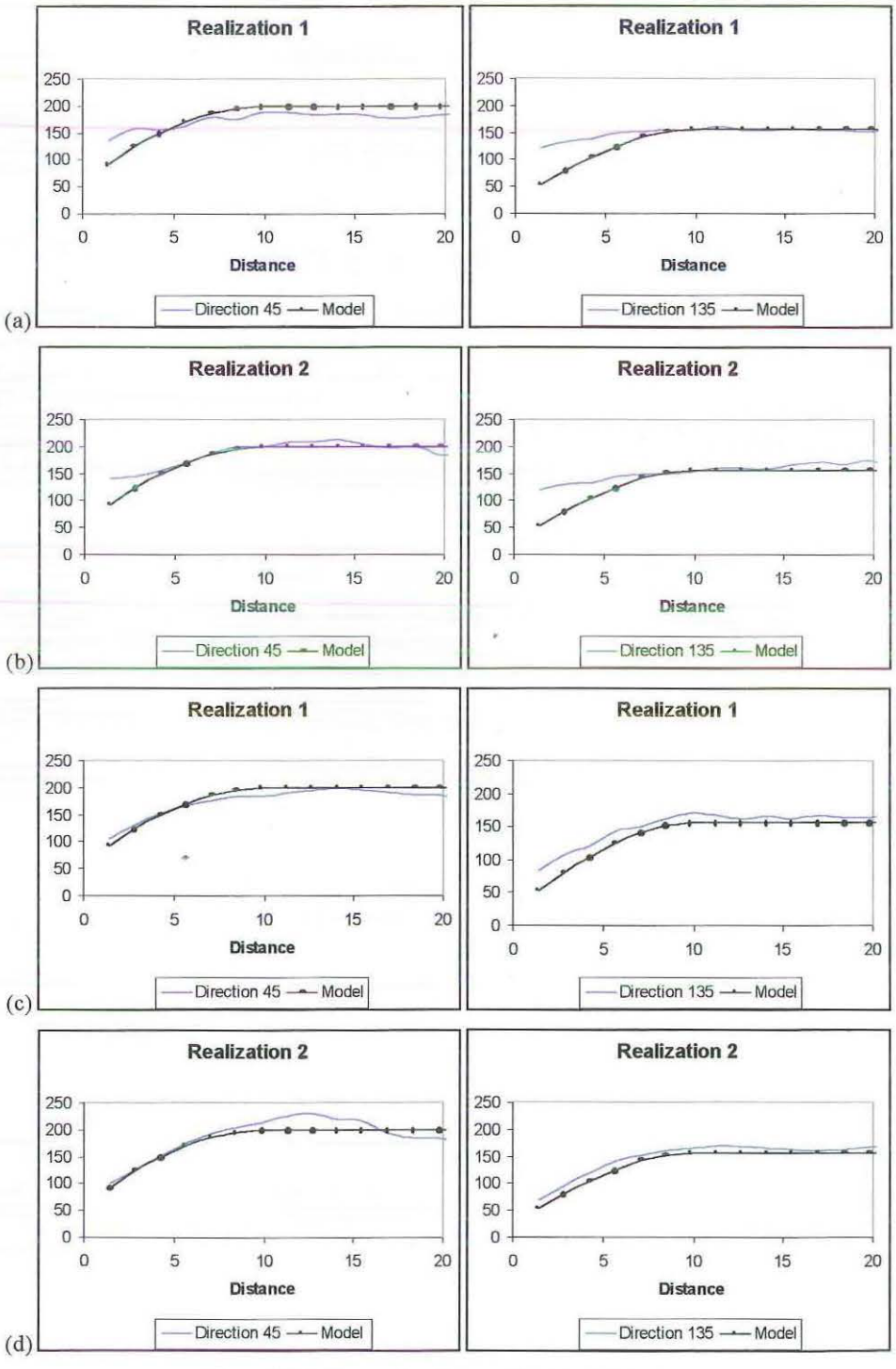

Figure 7.24: The semivariograms for the realizations.

The semivariograms for the realizations with a weight of one in Figure 7.24a and Figure 7.24b have a nugget value that is greater than that of the model but the sill has been reached and therefore these realizations where the weight of the semivariogram component is one are not an improvement on the realizations found in section 7.4 where the weight of the semivariogram component was five.

The other semivariograms of the realizations with a weight of three in Figure 7.24c and Figure 7.24d match the required semivariogram models in particular in the minor direction, but the semivariograms in section 7.4 (Figure 7.18), where the weight of the semivariogram component was five, were better reproduced.

7.6 Experimenting with the annealing schedule.

In Section 6.5 we introduced the annealing schedule. Deutsch and Cockerham (1994) and Goovaerts (1997) examined different values for the annealing schedule parameters to determine different annealing speeds. These are seen in the table 7.1 and are default, fast and very fast annealing schedules. These are the annealing schedules that will be used and compared. The parameter file *TrVH.par* that uses nlags equal to 98 in Section 7.4, is altered for each annealing schedule to get three more parameter files, *TrAD.par*, *TrAF.par* and *TrAVF.par* These parameter files can be seen in Appendix C.6. The default annealing schedule in Table 7.1 is different from the default annealing schedule used in SASIM when we set the parameter file to the automatic annealing setting. In this section, to distinguish between these default settings, we will refer to SASIM's default annealing schedule as *automatic*. The default annealing schedule used in the SASIM algorithm was seen earlier in Section 7.1.

Table 7.1: Recommended annealing schedules.

Schedule	T_0	λ	K_{max}	K_{accept}	S	O_{min}
Quench	0.0	0.00	-	-	-	0.001
Very fast	0.5	0.01	10	2	3	0.001
Fast	1.0	0.05	50	5	3	0.001
Default	1.0	0.10	100	10	3	0.001

The maps of the realizations for the different annealing schedules used are seen in Figure 7.25, together with the maps of the exhaustive and sample data sets and the realizations from section 7.4.

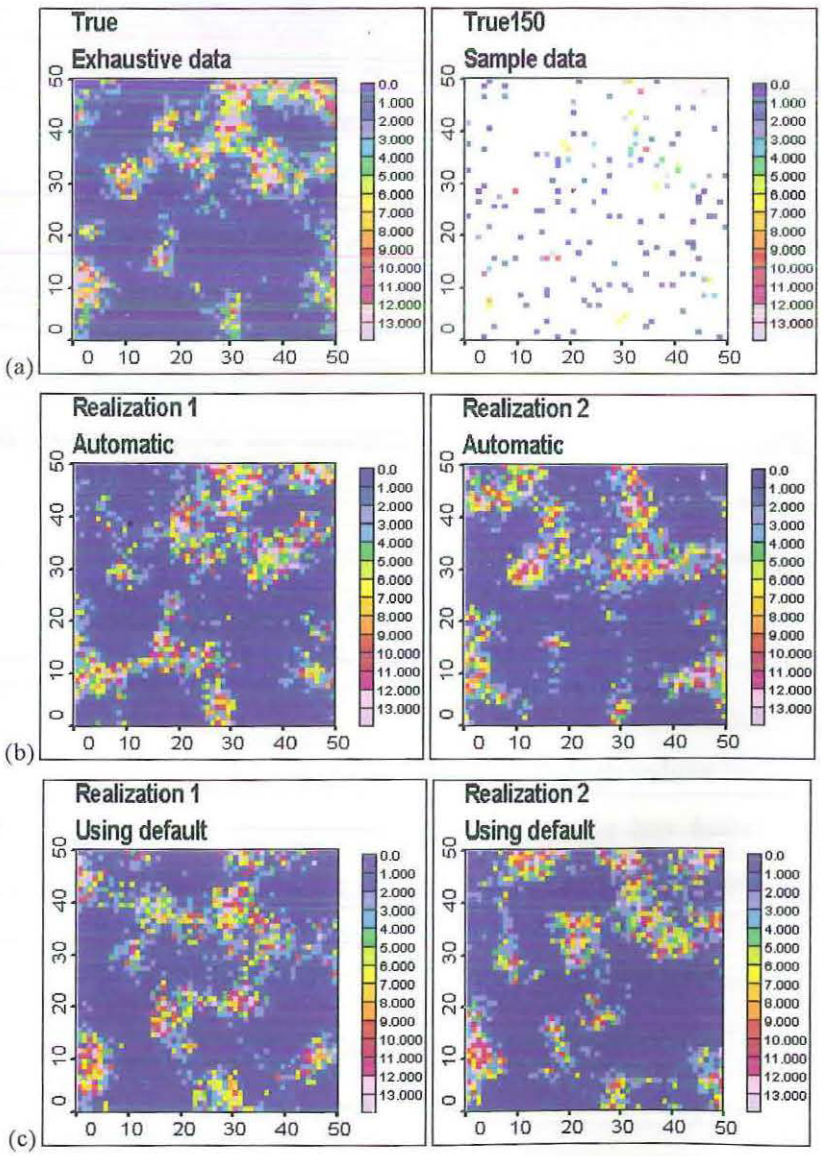

Figure 7.25: The maps of the realizations for the different annealing schedules.

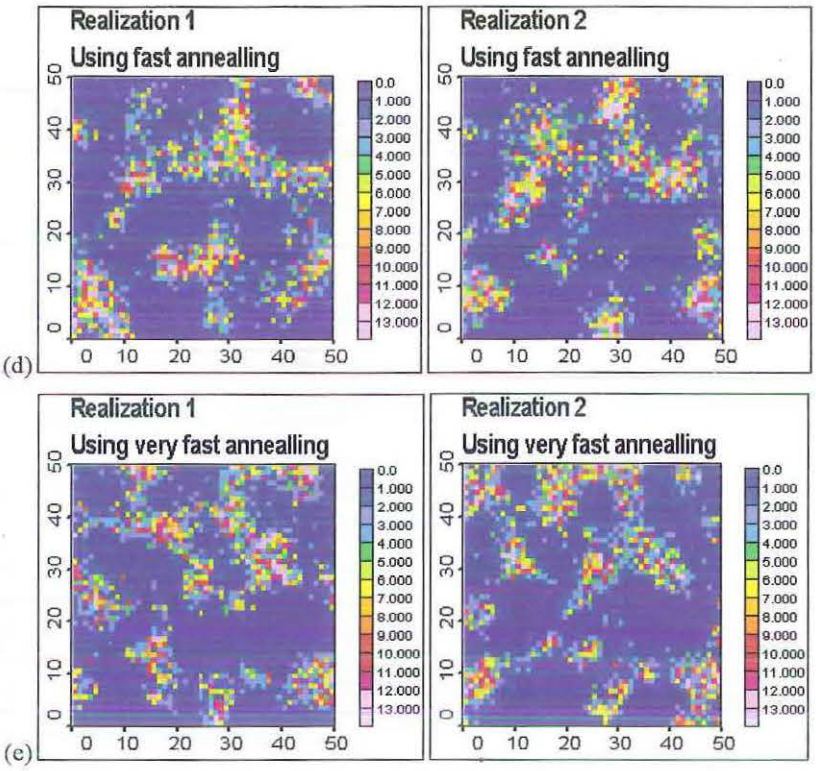

Figure 7.25continued.

The maps for the realizations in Figure 7.25d for the default annealing schedules have some similar aspects, but appear to have more high values in some regions especially around the region (30,20), than the map for the exhaustive data in (a). We have included the realizations for the SASIM automatic values in Figure 7.25c to see if there is any difference between these and the default annealing schedules from Deutsch and Cockerham (1994). There are some visible differences between the two different default annealing schedules.

Using fast and very fast annealing schedules the realizations in Figures 7.25e and 7.25f appear to have a higher number of high values in some regions, but fewer in others, such as the top right hand corner has fewer values. But generally the aspects seen in the map of the exhaustive data in (a) and the map of the sample data are seen in the realizations.

The histograms for the realizations for the different annealing schedules are in Figure 7.26, together with the histogram for the sample data set. The

histograms for the default, fast and very fast annealing schedules are in (b), (c) and (d) respectively.

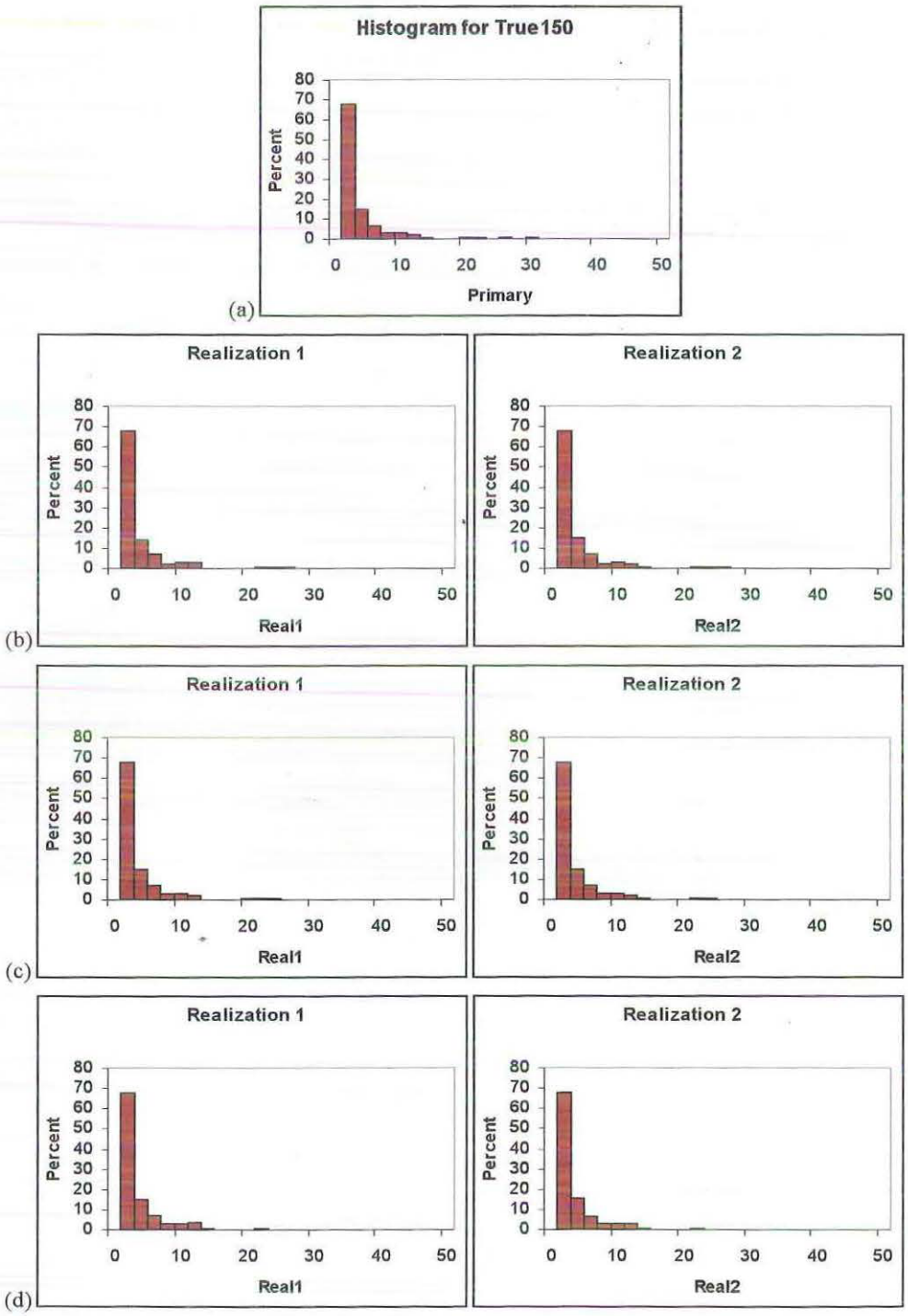

Figure 7.26: Histograms for the realizations.

The histograms for the realizations in Figure 7.26 when compared to the sample histogram in (a) are similar. That is, the histogram for the realizations in all cases has been reproduced from the histogram of the sample. They all begin

with a high number of values and taper off to approximately twelve, with several values to 30.

The omnidirectional semivariograms for the realizations of the different annealing schedules are in Figure 7.27. All contain the semivariogram model determined in section 3.3.

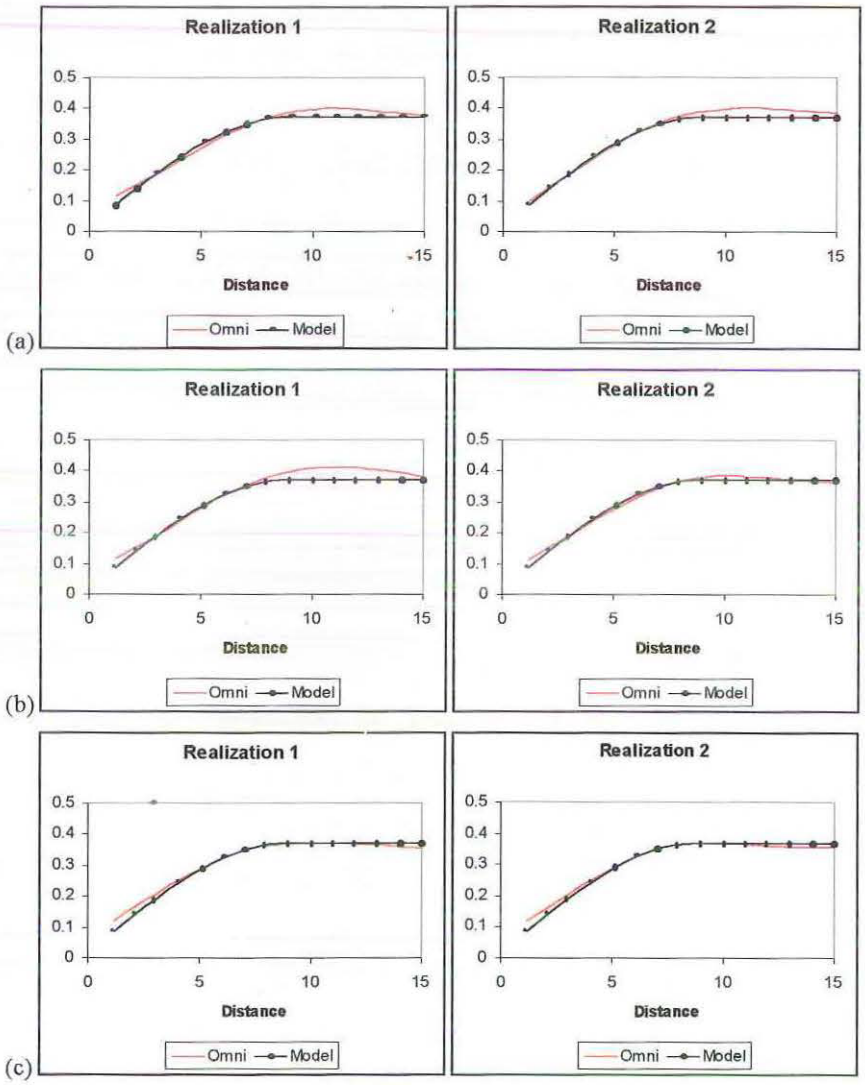

Figure 7.27: Semivariograms for the realizations for the different annealing schedules.

The omnidirectional semivariograms in Figure 7.27 are for the different annealing schedules. The semivariograms for the default (a) annealing schedule begin with a higher nugget effect than seen in the semivariogram model, and contains a greater range at approximately eleven. The semivariograms for the fast (b) and very fast (c) annealing schedules also contain a high nugget effect but the

range is similar to the semivariogram model, except for the first realization in (b). This shows that there are variations between the realizations.

7.7 Conclusions about the use of SASIM.

True150.

The first experiment in Section 7.3 sets only the semivariogram component to generate the realizations and to reproduce the omnidirectional semivariogram model. This showed that using nlags equal to 56 and nlags equal to 98 were too small since the semivariograms of the realizations did not reach the sill of the semivariogram model, and using the nlags equal to 120 it was seen that the sill was reached, but after the sill was reached, all omnidirectional semivariograms dipped to a lower value. The realizations were similar to the map of the exhaustive data, and the histograms of the realizations were different for the values between zero and one, but were then similar to the histogram of the sample data. This is a direct consequence of the sample histogram being used to generate the initial image.

The second experiment in Section 7.4 set the semivariogram and histogram components to attempt to reproduce the semivariogram model and histogram of the sample. Also the weight of the semivariogram component was set to five. It showed that using nlags equal to 56 still did not reach the sill, and nlags equal to 98 and 120 reached the sill, and the results were similar. But using nlags equal to 98 was selected because it used less time to generate the realizations. In all cases, including nlags equal to 56, the histograms from the realizations were successfully reproduced from the histogram of the sample data. Therefore setting the histogram component improves the results only slightly.

The third experiment in Section 7.5 used the histogram and semivariogram components, but changed the weight of the semivariogram. The results for the weight set to five were the best when compared with the results of changing the weight to one and three. As expected, the histograms were reproduced successfully, but there were differences in the semivariograms for the weights

equal to one and three. All semivariograms initially began with higher nugget values and a longer range than the model.

So from experimenting with and without the histogram component set, and using different weights, it is seen that using simulated annealing with *True150* the best reproduction of the histograms and semivariograms was determined by setting histogram and semivariogram components in the objective function, as well as setting the weight of the semivariogram component to five. We saw the resulting omnidirectional experimental semivariograms for two realizations and it is useful to find more, to check the exactness of the semivariogram reproduction. This is in Figure 7.28 with the model included.

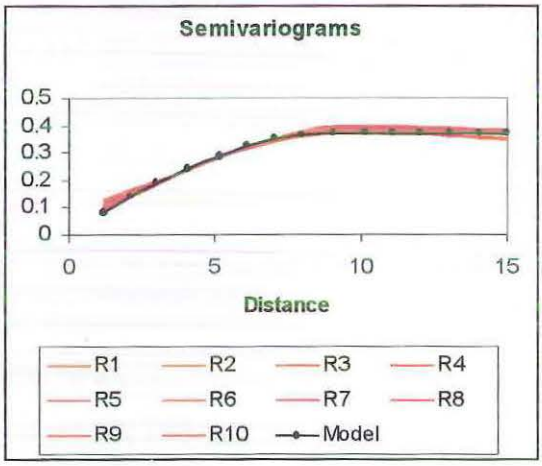

Figure 7.28: Omnidirectional semivariograms for the first ten realizations.

The omnidirectional semivariograms in Figure 7.28 show that there are small differences, such as a higher nugget effect and some fall below the semivariogram model, but we can see that the omnidirectional semivariogram model has been reproduced.

The last experiment in Section 7.6 used three different annealing schedules, the default, fast and very fast. It is seen that the default annealing schedule for SASIM and from Deutsch and Cockerham (1994) show some differences. But generally all realizations showed similar aspects to the exhaustive and sample data seen through the maps, the histograms and the omnidirectional semivariograms. The histograms of the realizations reproduced

the histogram of the sample, and the semivariograms contained higher nugget effects and some have greater ranges, but the general pattern is followed. The best realizations were still from Section 7.4, where the semivariogram and histogram components were set in the objective function.

Bereal00.

In Section 7.3, the first experiment used a parameter file that set only the semivariogram component. For all realizations with the value of the nlags as 80, 153 and 216, the result showed that the semivariogram in the direction 135° (the direction of maximum continuity) were almost perfectly reproduced, but the semivariogram in the perpendicular direction was not reproduced very well. For nlags equal to 80 and 153 the semivariograms in the direction 45° did not reach the sill of the required model. The histograms for the realizations, although not reproduced, were used by SASIM to generate the realizations.

The next experiment in Section 7.4 used a SASIM parameter file that set the histogram and the semivariogram components in the objective function, and the weight of the semivariogram component was set to five. Also, the nlags parameter used the values 80, 153 and 216. The histograms for the realizations were similar to the histogram of the sample. The semivariograms in direction 45° were better reproduced than the semivariograms in direction 45° in Section 7.3, and the semivariograms in direction 135° were worse than the semivariograms in direction 135° in Section 7.3. Therefore the inclusion of the histogram was necessary to gain a better reproduction of the semivariogram in the minor direction, and both semivariograms for each realization in Section 7.4 together are better than the results from just using the semivariogram component in Section 7.3.

Section 7.5 experimented with the weight of the semivariogram component, changing from five as in Section 7.4, to one and three. Also, the semivariogram and histogram components were still set. The results for Section 7.4 were reasonable. The results of the realizations with the weight equal to one showed that the histograms were similar to the histogram of the sample, but the

semivariograms were worse than the realizations with the weight equal to five. The results for the weight equal to three showed that the histograms for the realizations were similar to the histogram of the sample, and the semivariograms for the realizations were reasonable. But the semivariograms in Section 7.4 for nlags equal to 153 were better. The semivariograms in directions 45° and 135° for ten realizations are in Figure 7.69.

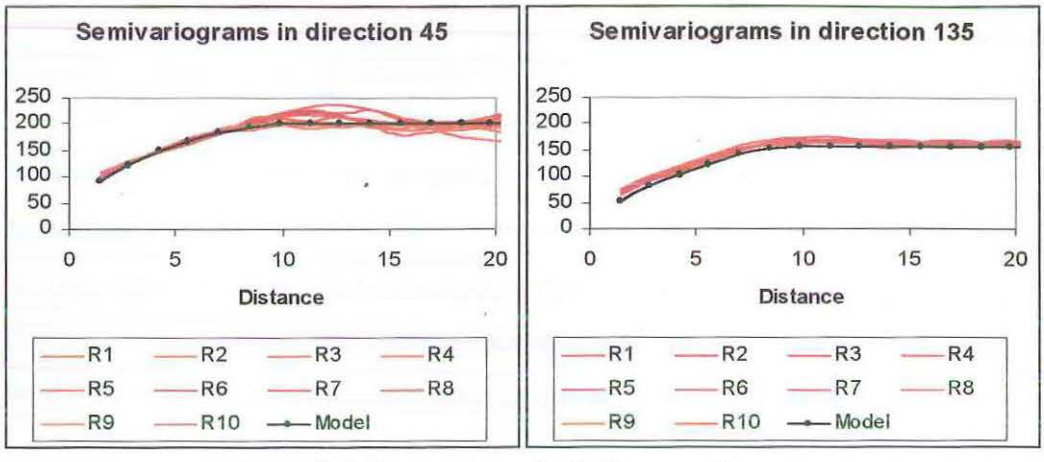

Figure 7.69: Semivariograms for the first ten realizations.

The semivariograms in the direction 135° have been reproduced, but the majority of the semivariograms are above the semivariogram model. The semivariograms in direction 45° although, have more variations than seen in direction 135°. Up to the lag distance of ten, the semivariograms are similar, but when the lag distance was greater than ten there were differences. Some are close to the semivariogram model, but others either tend off to a higher value or lower value.

8. COMPARISONS.

8.1 Comparison Methods.

In each method of simulation, we concentrated on finding a set of realizations that best represented the data set. For Sequential Gaussian simulation the parameter file was set from the cumulative distribution function of the sample data set. There was little experimenting that could be done with this parameter file. With Simulated Annealing, there were several parameters that we could experiment with. For this parameter file we experimented by using different annealing schedules (default, fast and very fast), by using one or two components in the objective function, by using different weights for the semivariogram component when we used two components, and by using different numbers of lags. We saw that the best set of realizations for both data sets used the histogram and semivariogram components in our objective function with the weight of the semivariogram component set to the maximum of five. Furthermore the number of lags selected was the middle value (for *True150* this was 98 and for *Berea100* it was 153), generally to save processing time and they reproduced the target constraints. That is, these realizations were required to reproduce the histogram and semivariograms of the sample.

In this section, we will compare the "best" realizations for Simulated Annealing with those from Sequential Gaussian simulation. To compare we will find the mean square error and the mean absolute deviation for each realization for each simulation method for each data set. We will start with mean square error followed by the mean absolute deviation.

The mean square error finds the average of the error squared, where the error is the difference at each location between the exhaustive set and the realization. That is, the mean square error for n observations is equal to:

$$\frac{\sum_{i=1}^{n}\left(z_s\left(\mathbf{u}_i'\right)-z\left(\mathbf{u}_i'\right)\right)^2}{n} \qquad (8.1)$$

The mean absolute deviation finds the average of the absolute error. That is the mean absolute deviation of n observations is computed as:

$$\frac{\sum_{i=1}^{n}\left|z_s\left(\mathbf{u}_i'\right)-z\left(\mathbf{u}_i'\right)\right|}{n} \tag{8.2}$$

8.2 Mean Square Error.

The first set of mean square error results in Figure 8.1 and Figure 8.2 are for *True150*. The results of the mean square error for 50 realizations generated by Sequential Gaussian simulation for *True150* are shown in Figure 8.1.

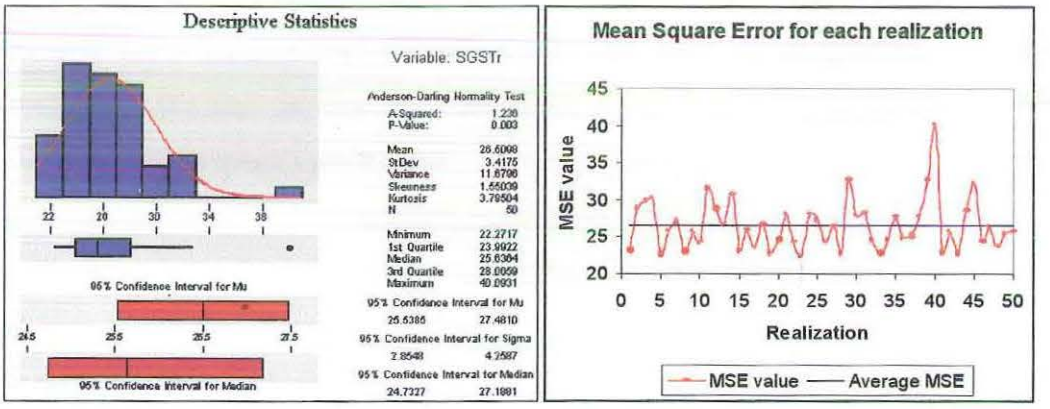

Figure 8.1: The mean square error for Sequential Gaussian Simulation on True150.

In Figure 8.1 we see that the average mean square error is approximately 26.51. The maximum is 40.09 from realization 40 and the minimum is 22.27 from realization 23, so the range of the mean square error is 17.82. The realization with the minimum mean square error can be considered the best representation of *True* while the maximum mean square is the worst. This range appears to be quite large, but if we consider that there is only one outlier, removing it would give us a much smaller range of approximately 13. There are also three realizations that have the average mean square error.

The results of the mean square error of 50 realizations for *True150* generated by Simulated Annealing are shown in Figure 8.2.

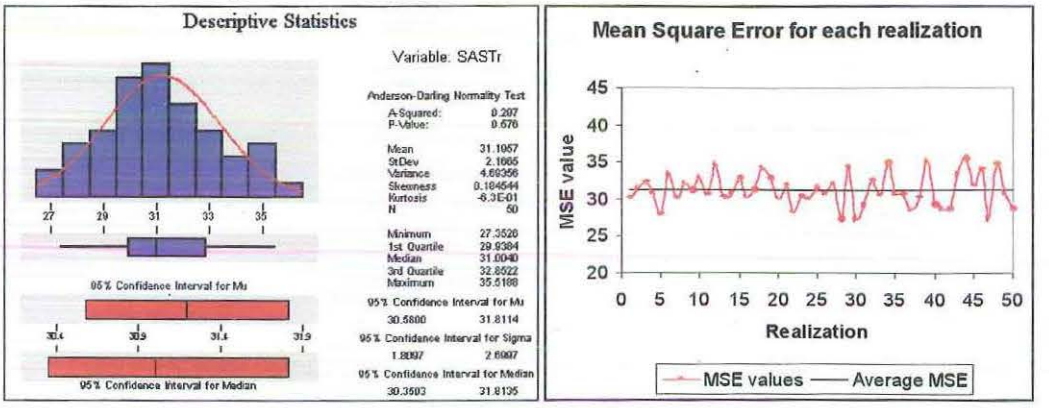

Figure 8.2: The mean square error for Simulated Annealing on True150.

In Figure 8.2, using Simulated Annealing we see that the average mean square error is 31.20, which is larger than the average of 26.51 seen for Sequential Gaussian simulation. The maximum of 35.52 is at realization 44 and the minimum of 27.35 is at realization 47. The range of mean square error is 8.17, which is much smaller than the range of 17.82 seen for Sequential Gaussian simulation. We also see that there are about ten realizations that are close to or have the average mean square error.

It appears that calculating and using the mean square error to compare between Sequential Gaussian and Simulated Annealing using *True150* that on average the best realizations are found using Sequential Gaussian simulation. That is the average mean square error is lower for Sequential Gaussian simulation and the minimum is also lower. But for Simulated Annealing, the range is much smaller and contains no outliers.

The results for the mean square error for Sequential Gaussian simulation for 50 realizations using *Berea100* are shown in Figure 8.3.

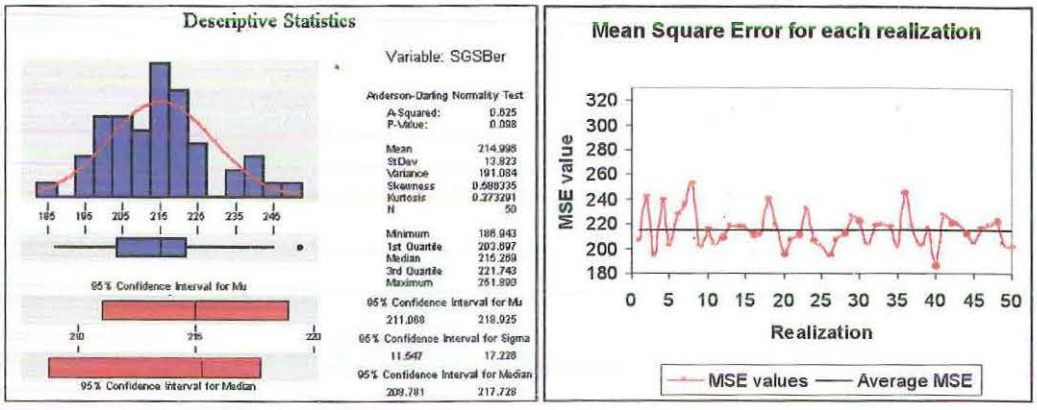

Figure 8.3: The mean square error for Sequential Gaussian Simulation on Berea100.

The average mean square error is approximately 215. The minimum of 185.94 is at realization 40 and the maximum of 251.89 is at realization 8, so the range is 65.95. The maximum is also an outlier. Furthermore, there are 13 realizations that are close to or equal the average mean square error.

The results for the 50 realizations for Simulated Annealing using *Berea100* are shown in Figure 8.4.

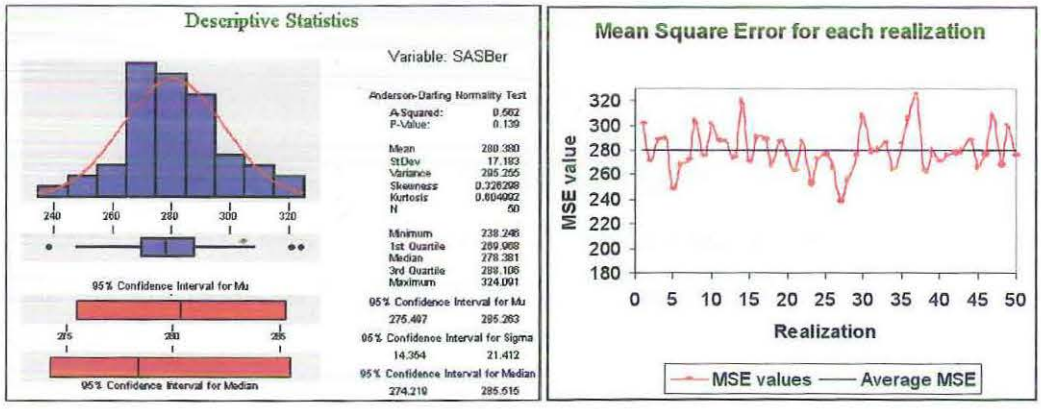

Figure 8.4: The mean square error for Simulated Annealing on Berea100.

In Figure 8.4 we see that the average mean square error is approximately 280.38 which is considerably higher than the average of 215 seen for Sequential Gaussian simulation. The minimum of 238.25 is at realization 27 and the maximum of 324.09 is at realization 37, so the range is 85.84. These are also much greater than the results for Sequential Gaussian simulation. We also see that the minimum and maximum are two of the three outliers found in Figure 8.4. There are also about 10 realizations that are close to, or equal to, the average mean

square error, which is less than the number found for Sequential Gaussian simulation.

Based on the examination of the mean square errors for the realizations generated by Sequential Gaussian and Simulated Annealing for *Berea100*, that the best results are determined by using Sequential Gaussian simulation. That is the average mean square error is lower, the range is smaller and there are fewer outliers found so the realizations are less erratic.

Also, if we examine the results gained from both data sets we see that the mean square error for *Berea100* is much larger than the mean square error of *True150*. This is because the values of the observations are much larger and there is a large variation of each value at each observation. That is, the mean, minimum and maximum for *Berea100* are 56.785, 25.000 and 89.000 respectively. When we compare these statistics with those from *True150* (mean equal to 2.600, minimum equal to 0.030 and maximum equal to 29.850), there is a large difference.

8.3 Mean Absolute Deviation.

In Figure 8.5 and Figure 8.6, we use the 50 realizations for *True150* generated with both simulation methods to find the mean absolute deviation. The mean absolute deviation for the realizations for Sequential Gaussian simulation is shown in Figure 8.5.

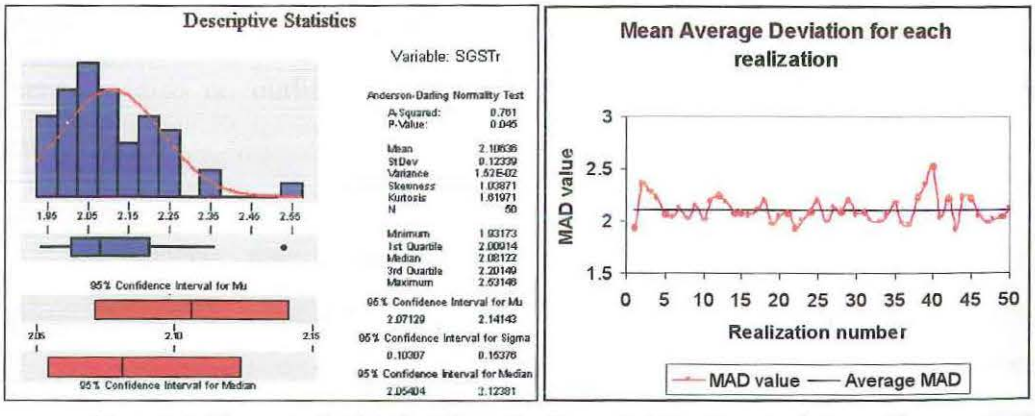

Figure 8.5: The mean absolute deviation using Sequential Gaussian simulation on True.

From Figure 8.5, we see that the average mean absolute deviation is approximately 2.11. The minimum of 1.93 is for realization 22 and the maximum of 2.53 is for realization 40, which gives us a range of 0.60. But this range could be smaller if we remove the one outlier seen in Figure 8.5a. There are also 14 realizations that are either equal to, or close to, the average mean absolute deviation.

The results from the mean absolute deviation for realisations for *True150* generated by Simulated Annealing are shown in Figure 8.6.

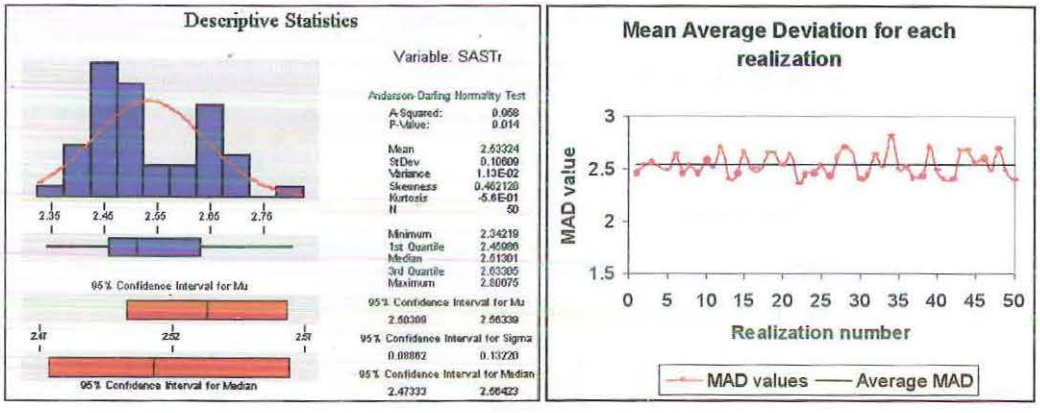

Figure 8.6: The mean absolute deviation using Simulated Annealing on True.

The average mean absolute deviation in Figure 8.6 for Simulated Annealing using *True150* is 2.53, which is greater than the average mean deviation seen for Sequential Gaussian simulation of 2.11. The minimum of 2.34 is at realization 22 and the maximum of 2.81 is at realization 34 and so the range is 0.47, which is smaller than the range found for Sequential Gaussian of 0.60. There are also no outliers found so we can assume that there are no erratic realizations.

It appears that using the mean absolute deviation to compare the realizations from Simulated Annealing and Sequential Gaussian simulation indicates the realizations for Simulated Annealing has a larger average mean absolute deviation and a higher minimum value. Also the range is smaller and the

data consist of no outliers. Therefore using Sequential Gaussian produces closer results to the exhaustive data. Using the mean squared error also indicated this result.

The next pair of results in Figure 8.7 and Figure 8.8 use *Berea100* to create 50 realizations each for Sequential Gaussian and Simulated Annealing. The results for Sequential Gaussian simulation are shown in Figure 8.7.

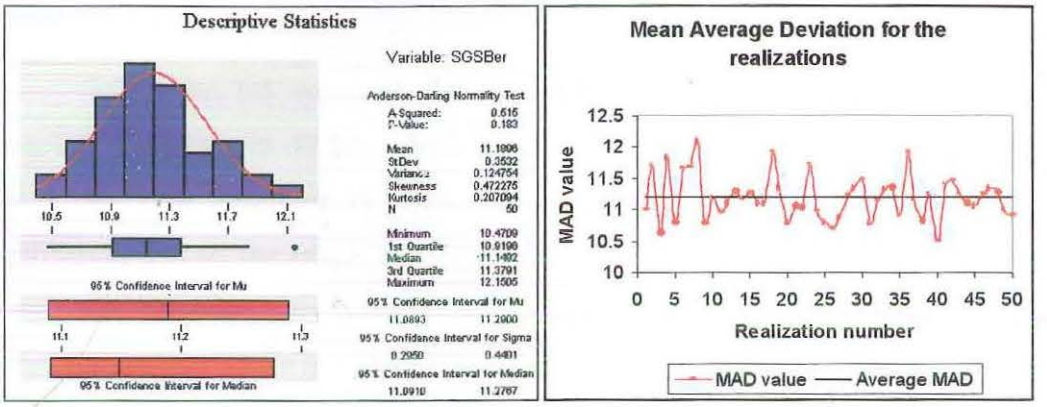

Figure 8.7: The mean absolute deviation using Sequential Gaussian simulation on Berea.

In Figure 8.7, the average mean absolute deviation for *Berea100* using Sequential Gaussian simulation is about 11.19. The minimum is 10.47 at realization 40 and the maximum is 12.15 at realization 8 so the range is 1.68. There is also one outlier, so the range could be smaller if we remove this outlier. There are also 10 realizations that have a mean absolute deviation that is close to, or equal to, the average.

The results for the mean absolute deviation for the 50 realizations using *Berea100* generated by Simulated Annealing are shown in Figure 8.8.

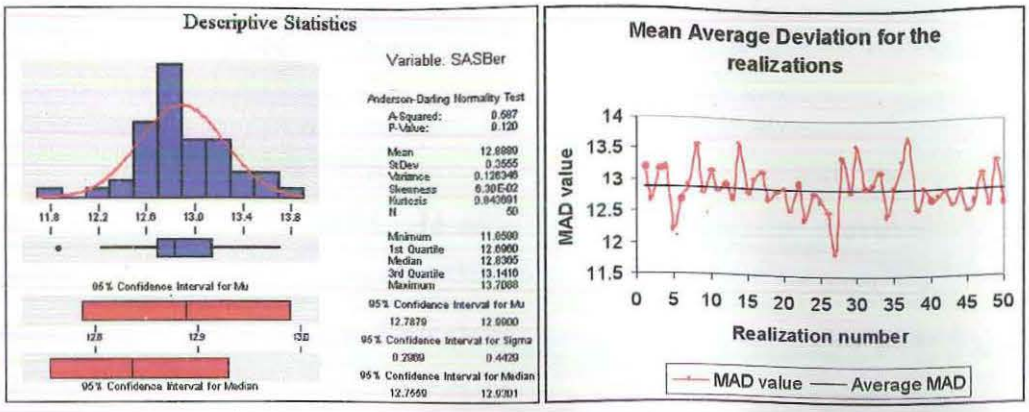

Figure 8.8: The mean absolute deviation using Simulated Annealing on Berea.

In Figure 8.8 we have the results for *Berea100*. The average mean absolute deviation is 12.89, which is greater than the average for Sequential Gaussian. The minimum is 11.86 at realization 27 and maximum is 13.71 at realization 37, so the range is 1.85, which is greater than the range found for Sequential Gaussian simulation. There is also one outlier found which is at the minimum value. There are also more realizations that are close to, or equal to, the average.

The realizations generated by Sequential Gaussian simulation for *Berea100* shows that the average mean absolute deviation is much lower than that from the realizations from Simulated Annealing. Also the minimum mean absolute deviation is much lower for Sequential Gaussian than for Simulated Annealing.

As mentioned earlier there is a difference between the values for the mean absolute deviation for *True150* and *Berea100* because of the larger variation for each observation in *Berea100*.

9. CONCLUSION.

This thesis has been devoted to experimentation with the Simulated Annealing algorithm SASIM. In addition, a comparison is provided with the Sequential Gaussian algorithm SGSIM. The methods were evaluated by means of the criteria of Mean Square Error (MSE) and Mean Absolute Deviation (MAD). As we saw, using Simulated Annealing is computer time intensive but, for the case studies we considered, did not produce a better result than that obtained from Sequential Gaussian simulation.

In the context of two case studies involving both sample and exhaustive data sets, we demonstrated the use of SASIM and then investigated the effect on the outcome of varying the different algorithm parameters, such as number of lags, and also the annealing schedule. We also considered the effect of changing the weighting given in the simulated annealing objective function to the reproduction of each of the sample histogram and semivariogram. We saw that the best set of realizations for both case studies used the objective function with both components included and with the weight of the semivariogram component set to the maximum of five. Furthermore we found that a 'middle' choice for the number of lags (for *True150* this was 98 and for *Berea100* it was 153), to save processing time, still produced a reasonable result, with good reproduction of the target constraints.

In industry, there is always the need to fully understand the various methods on offer and to make an informed choice between them in any particular case. This study is useful from this point of view.

REFERENCES.

Aarts, E.H. (1989) *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimisation and Neural Computing*, Wiley, John and Sons, Inc.

Armstrong, M. and Dowd, P.A (1994) *Geostatistical Simulation*, Kluwer Academic Publishers, Netherlands.

Azencott, R. (1991) *Simulated Annealing: Parallelization Techniques*, Wiley, John and Sons, Inc.

Caers, J. (2001) *Automatic Histogram and Variogram Reproduction in Simulated Annealing Simulation*, Mathematical Geology, v33 n2, pp 167-190, Kluwer Academic/Plenum Press.

Carle, S.F. (1997) *Implementation Schemes for Avoiding Artifact Discontinuities in Simulated Annealing*, Mathematical Geology, v29 n2, p 231, Kluwer Academic/Plenum Press.

Chiles, J-P. and Delfiner, P. (1999) *Geostatistics: measuring spatial uncertainty*, Wiley, New York.

Chu, J., Xu, W., Zhu, H., and Journel, A.G. (1991) *The AMOCO case study*, In Report 4, Stanford Center for Reservoir Forecasting, Stanford, CA.

Davis, L. (1990) *Genetic Algorithms and Simulated Annealing*, Pitman Publishing, London.

Deutsch, C.V. and Cockerham, P.W. (1994) *Practical Considerations in the Application of Simulated Annealing to Stochastic Simulation*, Mathematical Geology v26 no1, pp 67-82, Kluwer Academic/Plenum Press.

Deutsch, C.V. and Journel, A.G. (1998) *GSLIB Geostatistical Software Library and User's Guide*, Oxford University Press, New York.

Deutsch, C.V. and Wen, X.H. (1998) *An Improved Perturbation Mechanism for Simulated Annealing Simulation*, Mathematical Geology, v30 n7, p 801, Kluwer Academic/Plenum Press.

Fang, J.H. and Wang, P.P. (1997) *Random Field Generation Using Simulated Annealing vs. Fractal-Based Stochastic Interpolation*, Mathematical Geology, v29 n6, p 849, Kluwer Academic/Plenum Press.

Goovaerts, P. (1997) *Geostatistics for Natural Resources Evaluation*, Oxford University Press, New York.

Goovaerts, P. (1998) *Accounting for Estimation Optimality Criteria in Simulated Annealing*, Mathematical Geology, v30 n5, p 511, Kluwer Academic/Plenum Press.

Gotway, C.A. and Rutherford, B.M. (1994) *Stochastic Simulation for Imaging Spatial Uncertainty: Comparison and Evaluation of Available Algorithms*, in Armstrong M. and Dowd, P.A (editors), *Geostatistical Simulations*, pp 1-21, Kluwer Academic Publishers, Netherlands.

Hegstad, B.K., Omre, H., Tjelmeland, H. and Tyler, K. (1994) *Stochastic Simulation and Conditioning by Annealing in Reservoir Description*, in Armstrong, M. and Dowd, P.A (editors), *Geostatistical Simulations*, pp 1-21, Kluwer Academic Publishers, Netherlands.

Isaaks, E.H. and Srivastava, R.M. (1989) *An Introduction to Applied Geostatistics*, Oxford University Press, New York.

Kalivas, J.H. (1995) *Adaption of Simulated Annealing to Chemical Optimization Problems*, Elsevier Science.

Kirkpatrick, S., Gelatt, C.D. and Vecchi, M.P. (1983) *Optimisation by Simulated Annealing*, Science, v220, pp 671-680.

Li, P., Lowe, K. and Wozny, G. (2000) *Integration of Simulated Annealing to a Simulation Tool for Dynamic Optimization of Chemical Processes*, Chemical engineering and processing, v39 n4, p357.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and Teller, E. (1953) *Equation of State Calculations by Fast Computing Machines*, Journal of Chemical Physics, v21, p 1087-1090.

Nolte, A. and Schrader, R. (2001) *Simulated Annealing and Graph Colouring*, Combinatorics Probability and Computing, v10 n1, p29-40, Cambridge University Press.

Rivoirard, J., Simmonds, J., Foote, K.G., Fernandes, P. and Bez, N. (2000) *Geostatistics for Estimating Fish Abundance*, Blackwell Sciences Ltd, Great Britain.

Rosenhouse-Dantsker, A. and Osman, R. (2000) *Application of the Primary Hydration Shell Approach to Locally Enhanced Sampling Simulated Annealing: Computer Simulation of Thyrotropin-Releasing Hormone in Water*, Biophysical Journal, v79 n1, p66, Rockefeller University Press.

Wang, P., Dueck, G.W. and MacMillan, S. (2001) *Using simulated annealing to construct extremal graphs*, Discrete Mathematics, v235, p125-135, Elsevier Science Publishers.

APPENDIX A: THE DATA SETS.

A.1 *True150.*

x	y	Primary
44.5	1.5	0.04
3.5	36.5	0.24
27.5	23.5	0.82
46.5	31.5	2.38
5.5	15.5	0.15
22.5	39.5	0.86
35.5	38.5	0.93
41.5	12.5	0.68
22.5	7.5	0.28
37.5	23.5	0.12
15.5	15.5	10.82
41.5	5.5	0.78
13.5	24.5	0.84
16.5	30.5	0.43
23.5	6.5	1.50
40.5	39.5	0.37
19.5	34.5	3.68
17.5	33.5	0.93
13.5	1.5	0.11
4.5	49.5	1.19
25.5	6.5	1.15
5.5	2.5	0.05
3.5	48.5	0.63
47.5	23.5	1.18
2.5	19.5	2.55
10.5	8.5	0.34
36.5	29.5	6.61
31.5	35.5	3.08
0.5	41.5	2.33
33.5	34.5	1.88
38.5	31.5	18.64
15.5	3.5	0.10
4.5	13.5	3.88
36.5	33.5	2.38
3.5	12.5	11.50
17.5	28.5	0.17
21.5	49.5	0.05
44.5	15.5	0.48
46.5	23.5	0.64
22.5	30.5	1.21
2.5	28.5	0.31
47.5	40.5	1.31
32.5	40.5	5.62
46.5	11.5	5.65
17.5	26.5	0.39
18.5	37.5	5.25
43.5	18.5	0.63
37.5	10.5	0.77
31.5	36.5	4.29
10.5	19.5	0.03

x	y	Primary
3.5	46.5	0.72
45.5	28.5	0.71
14.5	24.5	0.28
33.5	47.5	9.46
16.5	16.5	29.85
43.5	12.5	1.00
32.5	39.5	3.84
30.5	42.5	21.08
3.5	0.5	0.95
36.5	49.5	1.96
7.5	48.5	1.89
43.5	29.5	3.59
27.5	34.5	1.82
4.5	6.5	6.70
14.5	15.5	2.36
9.5	28.5	9.27
19.5	18.5	1.30
31.5	33.5	2.79
23.5	18.5	0.16
20.5	39.5	0.57
19.5	36.5	7.21
33.5	22.5	0.78
9.5	15.5	0.08
39.5	10.5	0.29
11.5	6.5	0.24
45.5	11.5	2.11
14.5	20.5	0.06
13.5	35.5	1.19
39.5	16.5	0.08
29.5	3.5	6.01
28.5	14.5	0.11
34.5	16.5	0.07
42.5	18.5	0.56
0.5	32.5	0.87
40.5	32.5	4.66
44.5	26.5	0.81
7.5	25.5	0.17
20.5	11.5	0.99
35.5	3.5	0.33
15.5	8.5	0.60
27.5	10.5	0.30
20.5	45.5	2.03
27.5	46.5	1.81
45.5	8.5	2.31
49.5	31.5	1.27
0.5	23.5	0.66
43.5	22.5	0.69
4.5	21.5	1.75
48.5	0.5	0.95
49.5	21.5	1.68

x	y	Primary
5.5	25.5	0.12
31.5	43.5	5.39
4.5	7.5	8.56
40.5	44.5	1.10
20.5	43.5	0.82
3.5	29.5	0.32
2.5	26.5	0.09
43.5	17.5	0.25
37.5	34.5	4.09
26.5	28.5	1.18
34.5	36.5	11.42
29.5	23.5	1.14
39.5	37.5	1.66
45.5	47.5	24.03
48.5	26.5	0.63
20.5	1.5	0.03
30.5	10.5	0.63
20.5	24.5	1.74
43.5	44.5	1.30
40.5	33.5	5.49
17.5	0.5	0.82
47.5	8.5	3.92
44.5	27.5	0.91
0.5	37.5	0.56
30.5	49.5	12.08
24.5	10.5	0.94
27.5	31.5	1.87
24.5	33.5	2.88
24.5	26.5	0.58
30.5	4.5	5.51
7.5	14.5	0.39
6.5	27.5	2.53
34.5	42.5	2.20
4.5	40.5	0.78
17.5	34.5	1.36
36.5	17.5	0.10
33.5	18.5	0.37
15.5	9.5	0.25
32.5	37.5	3.21
25.5	9.5	0.79
31.5	9.5	1.57
34.5	12.5	0.46
7.5	46.5	1.03
21.5	16.5	0.25
34.5	2.5	0.11
38.5	29.5	2.68
24.5	37.5	3.90
17.5	15.5	9.08
31.5	49.5	5.49
15.5	35.5	2.18

A.2 *Berea100.*

Easting	Northing	Permeability
21.5	18.5	55.0
30.5	27.5	49.0
37.5	8.5	78.0
3.5	36.5	62.0
4.5	3.5	58.0
0.5	11.5	57.5
32.5	8.5	62.0
37.5	25.5	88.5
15.5	33.5	48.0
0.5	23.5	46.0
11.5	1.5	56.0
24.5	21.5	45.0
18.5	18.5	63.0
31.5	12.5	78.0
7.5	36.5	56.0
1.5	3.5	42.5
36.5	30.5	80.0
9.5	12.5	34.0
25.5	0.5	42.5
32.5	24.5	32.0
7.5	34.5	60.0
37.5	38.5	78.5
38.5	25.5	63.0
21.5	2.5	59.0
36.5	22.5	49.0
26.5	13.5	75.5
18.5	21.5	60.0
30.5	6.5	58.0
30.5	29.5	60.5
38.5	18.5	53.0
35.5	14.5	66.0
11.5	30.5	54.0
38.5	16.5	64.0
4.5	21.5	40.5
34.5	10.5	69.0
30.5	9.5	57.5
30.5	32.5	75.0
29.5	29.5	55.0
19.5	21.5	60.0
28.5	12.5	70.0
34.5	12.5	62.0
32.5	23.5	38.0
32.5	17.5	58.0
13.5	6.5	65.5
33.5	13.5	78.5
38.5	9.5	65.0
36.5	23.5	49.5
4.5	5.5	42.5
37.5	1.5	74.0
28.5	24.5	37.0

Easting	Northing	Permeability
21.5	4.5	42.0
23.5	28.5	50.0
0.5	19.5	39.0
24.5	8.5	58.5
38.5	37.5	82.0
30.5	10.5	58.5
24.5	13.5	75.0
21.5	26.5	56.0
10.5	12.5	36.0
6.5	21.5	30.0
8.5	36.5	55.0
37.5	35.5	89.0
28.5	8.5	54.0
35.5	28.5	67.0
20.5	7.5	32.5
31.5	10.5	67.0
18.5	4.5	54.0
22.5	10.5	51.0
21.5	3.5	38.5
35.5	22.5	40.0
1.5	22.5	65.0
39.5	31.5	77.5
15.5	24.5	55.5
27.5	25.5	50.0
5.5	39.5	53.0
24.5	36.5	79.5
5.5	5.5	58.5
18.5	33.5	55.0
30.5	37.5	89.0
18.5	34.5	48.0
4.5	12.5	60.0
14.5	20.5	60.0
14.5	11.5	30.0
30.5	7.5	51.0
36.5	0.5	59.0
35.5	5.5	69.0
14.5	3.5	49.5
38.5	29.5	75.0
21.5	13.5	45.5
31.5	31.5	85.0
31.5	1.5	25.0
2.5	34.5	61.0
23.5	14.5	48.0
4.5	26.5	46.0
39.5	14.5	56.5
7.5	23.5	40.0
20.5	28.5	49.0
2.5	29.5	48.0
6.5	2.5	38.0
11.5	36.5	46.5

APPENDIX B.

B.1 GSLIB2 library.

GSLIB is a suite of geostatistical programs that have be accumulated and described in Deutsch and Journel (1998). It consists executive programs that use user-defined parameter files to perform different geostatistical functions. The outputs from GSLIB are new data sets and sometimes debug files. In our case the outputs required will be to assist in obtaining and examining the simulated data. The programs and parameter files that will be used are:

NSCORE.EXE, NSCORE.PAR:

This program is used to transform non-normal data to normally distributed data. This is used in sequential Gaussian simulation.

VARMAP.EXE, VARMAP.PAR:

This is used to draw semivariograms surface maps where a large number of observations are used.

GAM.EXE, GAM.PAR:

This is used to draw semivariograms in different directions where the data is anisotropic.

GAMV.EXE, GAMV.PAR:

This is used to draw omnidirectional semivariograms for isotropic data sets.

SGSIM.EXE, SGSIM.PAR:

This is GSLIB's program to use for sequential Gaussian simulation.

SASIM.EXE, SASIM.PAR:

This is GSLIB's program to use for simulated annealing.

The parameter files are the files that contain the parameters to perform the required task. Not all parameter files have been included in Appendix C because there are too many parameter files to include. For example, when a parameter file for a simulation is run, the output produced includes data for three realizations. So when creating semivariogram surfaces (VARMAP) and semivariograms

(GAM and GAMV) for each realization, there would be numerous amounts of each. Therefore, only examples of each will be included.

Also, the parameter files have a line number for the parameters included. This is so when referencing to a particular line, it is easier to refer by line number.

B.2 3Plot98.

3Plot is a public domain program that is part of the GeoStat Office Suite. It is used to plot the observations into either a post plots or mosaic maps. It is also used for mapping the semivariogram surface.

B.3 Variowin 2.2.

Variowin is a program that consists of several executable programs, to create semivariograms, semivariogram surfaces, madograms, covariance and correlograms, and model the semivariograms.

B.4 Microsoft Excel 2000.

Microsoft Excel is one program found in the Microsoft Office Suite. It is a spreadsheet program.

B.5 Minitab 13

Minitab is a comprehensive statistical package, but in this study, it is used for exploratory data analysis to find the summary or descriptive statistics, histograms, box plots, and other information common to all other statistical techniques. It is also useful for calculating the logarithmic data, and exporting the columns of data to a data file.

APPENDIX C: PARAMETER FILES.

C.1 NORMAL SCORES TRANSFORMATION (NSCORE).

True150.

```
                    Parameters for NSCORE
                    *********************

   START OF PARAMETERS:
1  true150.dat              \file with data
2  3   0                    \  columns for variable and weight
3  -1.0e21   1.0e21         \  trimming limits
4  0                        \1=transform according to specified ref. dist.
5  histsmth.out             \  file with reference dist.
6  1   2                    \  columns for variable and weight
7  primns.dat               \file for output
8  primns.trn               \file for output transformation table
```

Berea100.

```
                    Parameters for NSCORE
                    *********************

   START OF PARAMETERS:
1  berea100.dat             \file with data
2  4   0                    \  columns for variable and weight
3  -1.0e21   1.0e21         \  trimming limits
4  0                        \1=transform according to specified ref. dist.
5  histsmth.out             \  file with reference dist.
6  1   2                    \  columns for variable and weight
7  b100ns.dat               \file for output
8  b100ns.trn               \file for output transformation table
```

C.2 VARMAP.

True150.

```
                    Parameters for VARMAP
                    *********************

    START OF PARAMETERS:
 1  truens.dat                      -file with data
 2  1   5                           -   number of variables: column numbers
 3  -1.0e21      1.0e21             -   trimming limits
 4  1                               -1=regular grid, 0=scattered values
 5   50   50    1                  -if =1: nx,     ny,   nz
 6  1.0  1.0  1.0                   -       xsiz, ysiz, zsiz
 7  1   2   0                       -if =0: columns for x,y, z coordinates
 8  TrnsVM.dat                      -file for variogram output
 9   30    30     0                 -nxlag, nylag, nzlag
10  1.0   1.0   1.0                 -dxlag, dylag, dzlag
11  1                               -minimum number of pairs
12  1                               -standardize sill? (0=no, 1=yes)
13  1                               -number of variograms
14  1   1   1                       -tail, head, variogram type
```

Berea100.

```
                    Parameters for VARMAP
                    *********************

    START OF PARAMETERS:
 1  bereans.dat                     -file with data
 2  1   5                           -   number of variables: column numbers
 3  -1.0e21      1.0e21             -   trimming limits
 4  1                               -1=regular grid, 0=scattered values
 5   40   40    1                  -if =1: nx,     ny,   nz
 6  1.0  1.0  1.0                   -       xsiz, ysiz, zsiz
 7  1   2   0                       -if =0: columns for x,y, z coordinates
 8  BrnsVM.dat                      -file for variogram output
 9   25    25     0                 -nxlag, nylag, nzlag
10  1.0   1.0   1.0                 -dxlag, dylag, dzlag
11  1                               -minimum number of pairs
12  1                               -standardize sill? (0=no, 1=yes)
13  1                               -number of variograms
14  1   1   1                       -tail, head, variogram type
```

C.3 GAMV.

True150.

```
                    Parameters for GAMV
                    *******************

    START OF PARAMETERS:
 1  primns1.dat                          -file with data
 2  1   2   0                            -   columns for X, Y, Z coordinates
 3  1   4                                -   number of variables,col numbers
 4  -1.0e21      1.0e21                  -   trimming limits
 5  Varns1.dat                           -file for variogram output
 6  10                                   -number of lags
 7  2.50                                 -lag separation distance
 8  1.25                                 -lag tolerance
 9  1                                    -number of directions
10  0.0 90.0 50.0   0.0 90.0 50.0        -azm,atol,bandh,dip,dtol,bandv
11  1                                    -standardize sills? (0=no, 1=yes)
12  1                                    -number of variograms
13  1   1   1                            -tail var., head var., variogram type
```

C.4 GAM.

Berea100.

```
                    Parameters for GAM
                    ******************

   START OF PARAMETERS:
1  bl00ns3.dat                -file with data
2  1   4                      -   number of variables, column numbers
3  -1.0e21     1.0e21         -   trimming limits
4  GBrns3.dat                 -file for variogram output
5  1                          -grid or realization number
6  40   0.5   1.0             -nx, xmn, xsiz
7  40   0.5   1.0             -ny, ymn, ysiz
8    1   0.5   1.0             -nz, zmn, zsiz
9  2  25                      -number of directions, number of lags
10   1  1  0                   -ixd(1),iyd(1),izd(1)
11   1 -1  0                   -ixd(2),iyd(2),izd(2)
12 1                          -standardize sill? (0=no, 1=yes)
13 1                          -number of variograms
14 1   1   1                  -tail variable, head variable, variogram type
```

C.5 SEQUENTIAL GAUSSIAN SIMULATION (SGSIM).

True150.

```
                  Parameters for SGSIM
                  ********************

   START OF PARAMETERS:
 1 true150.dat                  -file with data
 2 1  2  0  3  0  0             -  columns for x,y,z,vr,wt,sec.var.
 3 -1.0e21    1.0e21            -  trimming limits
 4 1                            -transform the data (0=no, 1=yes)
 5 sgsim.trn                    -  file for output trans table
 6 0                            -  consider ref. dist (0=no, 1=yes)
 7 histsmth.out                 -  file with ref. dist distribution
 8 1  2                         -  columns for vr and wt
 9 0.0    50.0                  -  zmin,zmax(tail extrapolation)
10 1       0.0                  -  lower tail option, parameter
11 4       1.25                 -  upper tail option, parameter
12 2                           -debugging level: 0,1,2,3
13 truesgs.dbg                 -file for debugging output
14 truesgs.dat                 -file for simulation output
15 50                          -number of realizations to generate
16 50    0.5    1.0            -nx,xmn,xsiz
17 50    0.5    1.0            -ny,ymn,ysiz
18 1     0.5    1.0            -nz,zmn,zsiz
19 69069                       -random number seed
20 0     8                     -min and max original data for sim
21 12                          -number of simulated nodes to use
22 1                           -assign data to nodes (0=no, 1=yes)
23 0     3                     -multiple grid search (0=no, 1=yes),num
24 0                           -maximum data per octant (0=not used)
25 10.0  10.0  10.0            -maximum search radii (hmax,hmin,vert)
26   0.0   0.0   0.0            -angles for search ellipsoid
27 0      1.0   1.0            -ktype: 0=SK,1=OK,2=LVM,3=EXDR,4=COLC
28 ../data/ydata.dat           -  file with LVM, EXDR, or COLC variable
29 0                           -  column for secondary variable
30 1    0.01                   -nst, nugget effect
31 1    0.99  0.0   0.0   0.0 -it,cc,ang1,ang2,ang3
32              9.6   9.6   0.0 -a_hmax, a_hmin, a_vert
```

```
                  Parameters for SGSIM
                  ********************

   START OF PARAMETERS:
 1 berea100.dat                  -file with data
 2 1  2  0  4  0  0              -  columns for X,Y,Z,vr,wt,sec.var.
 3 -1.0e21    1.0e21             -  trimming limits
 4 1                             -transform the data (0=no, 1=yes)
 5 sgsim.trn                     -  file for output trans table
 6 0                             -  consider ref. dist (0=no, 1=yes)
 7 histsmth.out                  -  file with ref. dist distribution
 8 1  2                          -  columns for vr and wt
 9 20.0   95.0                   -  zmin,zmax(tail extrapolation)
10 2       2.0                  -  lower tail option, parameter
11 4       1.5                  -  upper tail option, parameter
12 2                            -debugging level: 0,1,2,3
13 bersgs.dbg                   -file for debugging output
14 bersgs.dat                   -file for simulation output
15 50                           -number of realizations to generate
16 40    0.5    1.0             -nx,xmn,xsiz
17 40    0.5    1.0             -ny,ymn,ysiz
18 1     0.5    1.0             -nz,zmn,zsiz
19 69069                        -random number seed
20 0     8                      -min and max original data for sim
21 12                           -number of simulated nodes to use
22 1                            -assign data to nodes (0=no, 1=yes)
23 0     3                      -multiple grid search (0=no, 1=yes),num
24 0                            -maximum data per octant (0=not used)
25 30.0  15.0  15.0             -maximum search radii (hmax,hmin,vert)
26 135.0  0.0   0.0             -angles for search ellipsoid
27 0      1.0   1.0             -ktype: 0=SK,1=OK,2=LVM,3=EXDR,4=COLC
28 ../data/ydata.dat            -  file with LVM, EXDR, or COLC variable
29 0                            -  column for secondary variable
30 2    0.15                    -nst, nugget effect
31 1    0.53 135.0  0.0  0.0    -it,cc,ang1,ang2,ang3
32              7.0  7.0  0.0    -a_hmax, a_hmin, a_vert
33 1    0.32 135.0  0.0  0.0    -it,cc,ang1,ang2,ang3
34          12700.0 12.7  0.0    -a_hmax, a_hmin, a_vert
```

C.6 SIMULATED ANNEALING SIMULATION (SASIM)

True150 – Using only the semivariogram component (*TrV.par*).

```
                    Parameters for SASIM
                    ********************

   START OF PARAMETERS:
 1 0  1  0  0  0                  -components: hist,varg,ivar,corr,cpdf
 2 1  5  1  1  1                  -weight:     hist,varg,ivar,corr,cpdf
 3 1                              -0=no transform, 1=log transform
 4 50                             -number of realizations
 5 50       0.5     1.0           -grid definition: nx,xmn,xsiz
 6 50       0.5     1.0           -                 ny,ymn,ysiz
 7  1       0.5     1.0           -                 nz,zmn,zsiz
 8 69069                          -random number seed
 9 2                              -debugging level
10 truesasv.dbg                   -file for debugging output
11 truesasv.dat                   -file for simulation output
12 0                              -schedule (0=automatic,1=set below)
13 1.0  0.10  100  10  3  0.001   -   schedule: t0,redfac,ka,k,num,Omin
14 100  1.0                       -   maximum perturbations, reporting
15 1000                           -   maximum number without a change
16 1                              -conditioning data:(0=no, 1=yes)
17 true150.dat                    -   file with data
18 1   2   0   3                  -   columns: x,y,z,attribute
19 -1.0e21     1.0e21             -   trimming limits
20 1                              -file with histogram:(0=no, 1=yes)
21 true150.dat                    -   file with histogram
22 3   0                          -   column for value and weight
23 48                             -   number of quantiles for obj. func.
24 1                              -number of indicator variograms
25 2.78                           -   indicator thresholds
26 ../data/seisdat.dat            -file with gridded secondary data
27 1                              -   column number
28 1                              -   vertical average (0=no, 1=yes)
29 0.60                           -correlation coefficient
30 ../data/cal.dat                -file with paired data
31 2    1    0                    -   columns for primary, secondary, wt
32 -0.5    100.0                  -   minimum and maximum
33 5                              -   number of primary thresholds
34 5                              -   number of secondary thresholds
35 120                            -Variograms: number of lags
36 0                              -   standardize sill (0=no,1=yes)
37 1    0.01                      -   nst, nugget effect
38 1    0.36   0.0   0.0  0.0     -   it,cc,ang1,ang2,ang3
39              8.75  8.75 0.0     -    a_hmax, a_hmin, a_vert
```

True150 – Using the histogram and semivariogram components (*TrVH.par*).

```
                  Parameters for SASIM
                  ********************

   START OF PARAMETERS:
 1 1  1   0   0   0                    -components: hist,varg,ivar,corr,cpdf
 2 1  5   1   1   1                    -weight:     hist,varg,ivar,corr,cpdf
 3 1                                   -0=no transform, 1=log transform
 4 50                                  -number of realizations
 5 50        0.5     1.0               -grid definition: nx,xmn,xsiz
 6 50        0.5     1.0               -                 ny,ymn,ysiz
 7  1        0.5     1.0               -                 nz,zmn,zsiz
 8 69069                               -random number seed
 9 2                                   -debugging level
10 TrVH.dbg                            -file for debugging output
11 TrVH.dat                            -file for simulation output
12 0                                   -schedule (0=automatic,1=set below)
13 1.0   0.10  100    10   3   0.001   -   schedule: t0,redfac,ka,k,num,Omin
14 100   1.0                           -   maximum perturbations, reporting
15 1000                                -   maximum number without a change
16 1                                   -conditioning data:(0=no, 1=yes)
17 true150.dat                         -   file with data
18 1    2    0    3                    -   columns: x,y,z,attribute
19 -1.0e21     1.0e21                  -   trimming limits
20 1                                   -file with histogram:(0=no, 1=yes)
21 true150.dat                         -   file with histogram
22 3    0                              -   column for value and weight
23 48                                  -   number of quantiles for obj. func.
24 1                                   -number of indicator variograms
25 2.78                                -   indicator thresholds
26 ../data/seisdat.dat                 -file with gridded secondary data
27 1                                   -   column number
28 1                                   -   vertical average (0=no, 1=yes)
29 0.60                                -correlation coefficient
30 ../data/cal.dat                     -file with paired data
31 2     1     0                       -   columns for primary, secondary, wt
32 -0.5     100.0                      -   minimum and maximum
33 5                                   -   number of primary thresholds
34 5                                   -   number of secondary thresholds
35 120                                 -Variograms: number of lags
36 0                                   -   standardize sill (0=no,1=yes)
37 1     0.01                          -   nst, nugget effect
38 1     0.36    0.0    0.0   0.0      -   it,cc,ang1,ang2,ang3
39               8.75   8.75  0.0      -   a_hmax, a_hmin, a_vert
```

Berea100 – Using only the semivariogram component (*BerV.par*).

```
                  Parameters for SASIM
                  ********************

   START OF PARAMETERS:
 1 0  1  0  0  0                  -components: hist,varg,ivar,corr,cpdf
 2 1  5  1  1  1                  -weight:     hist,varg,ivar,corr,cpdf
 3 0                              -0=no transform, 1=log transform
 4 50                             -number of realizations
 5 40       0.5     1.0           -grid definition: nx,xmn,xsiz
 6 40       0.5     1.0           -                 ny,ymn,ysiz
 7   1       0.5     1.0          -                 nz,zmn,zsiz
 8 69069                          -random number seed
 9 2                              -debugging level
10 BerV216.dbg                    -file for debugging output
11 BerV216.dat                    -file for simulation output
12 0                              -schedule (0=automatic,1=set below)
13 1.0  0.10  100   10  3  0.001  -   schedule: t0,redfac,ka,k,num,Omin
14 100  1.0                       -   maximum perturbations, reporting
15 1000                           -   maximum number without a change
16 1                              -conditioning data:(0=no, 1=yes)
17 berea100.dat                   -   file with data
18 1   2   0   4                 -   columns: x,y,z,attribute
19 -1.0e21    1.0e21              -   trimming limits
20 1                              -file with histogram:(0=no, 1=yes)
21 berea100.dat                   -   file with histogram
22 4   0                          -   column for value and weight
23 24                             -   number of quantiles for obj. func.
24 1                              -number of indicator variograms
25 2.78                           -   indicator thresholds
26 ../data/seisdat.dat            -file with gridded secondary data
27 1                              -   column number
28 1                              -   vertical average (0=no, 1=yes)
29 0.60                           -correlation coefficient
30 ../data/cal.dat                -file with paired data
31 2    1    0                   -   columns for primary, secondary, wt
32 -0.5    100.0                 -   minimum and maximum
33 5                              -   number of primary thresholds
34 5                              -   number of secondary thresholds
35 216                            -Variograms: number of lags
36 0                              -   standardize sill (0=no,1=yes)
37 2    25.5                      -   nst, nugget effect
38 1   130.2 135.0   0.0 0.0      -   it,cc,ang1,ang2,ang3
39              9.9   9.9 0.0     -    a_hmax, a_hmin, a_vert
40 1    44.1 135.0   0.0 0.0      -   it,cc,ang1,ang2,ang3
41           2100.0   2.1 0.0     -    a_hmax, a_hmin, a_vert
```

Berea100 – Using the histogram and semivariogram components (*BerVH.par*).

```
                  Parameters for SASIM
                  ********************

    START OF PARAMETERS:
 1  1  1  0  0  0                  -components: hist,varg,ivar,corr,cpdf
 2  1  5  1  1  1                  -weight:     hist,varg,ivar,corr,cpdf
 3  0                               -0=no transform, 1=log transform
 4  50                              -number of realizations
 5  40       0.5     1.0           -grid definition: nx,xmn,xsiz
 6  40       0.5     1.0           -                 ny,ymn,ysiz
 7   1       0.5     1.0           -                 nz,zmn,zsiz
 8  69069                           -random number seed
 9  2                               -debugging level
10  BerVH216.dbg                    -file for debugging output
11  BerVH216.dat                    -file for simulation output
12  0                               -schedule (0=automatic,1=set below)
13  1.0  0.10  100  10  3  0.001    -   schedule: t0,redfac,ka,k,num,Omin
14  100  1.0                       -   maximum perturbations, reporting
15  1000                            -   maximum number without a change
16  1                               -conditioning data:(0=no, 1=yes)
17  berea100.dat                    -   file with data
18  1   2   0   4                  -   columns: x,y,z,attribute
19  -1.0e21     1.0e21             -   trimming limits
20  1                               -file with histogram:(0=no, 1=yes)
21  berea100.dat                    -   file with histogram
22  4   0                           -   column for value and weight
23  24                              -   number of quantiles for obj. func.
24  1                               -number of indicator variograms
25  2.78                            -   indicator thresholds
26  ../data/seisdat.dat             -file with gridded secondary data
27  1                               -   column number
28  1                               -   vertical average (0=no, 1=yes)
29  0.60                            -correlation coefficient
30  ../data/cal.dat                 -file with paired data
31  2    1    0                    -   columns for primary, secondary, wt
32  -0.5    100.0                   -   minimum and maximum
33  5                               -   number of primary thresholds
34  5                               -   number of secondary thresholds
35  216                             -Variograms: number of lags
36  0                               -   standardize sill (0=no,1=yes)
37  2    25.5                       -   nst, nugget effect
38  1   130.2 135.0   0.0 0.0       -   it,cc,ang1,ang2,ang3
39              9.9   9.9 0.0       -   a_hmax, a_hmin, a_vert
40  1    44.1 135.0   0.0 0.0       -   it,cc,ang1,ang2,ang3
41           2100.0   2.1 0.0       -   a_hmax, a_hmin, a_vert
```

TrAD.par- Using a default annealing schedule.

```
                  Parameters for SASIM
                  ********************

   START OF PARAMETERS:
1  1  1   0   0   0                  -components: hist,varg,ivar,corr,cpdf
2  1  5   1   1   1                  -weight:     hist,varg,ivar,corr,cpdf
3  1                                 -0=no transform, 1=log transform
4  50                                -number of realizations
5  50       0.5     1.0              -grid definition: nx,xmn,xsiz
6  50       0.5     1.0              -                 ny,ymn,ysiz
7   1       0.5     1.0              -                 nz,zmn,zsiz
8  69069                             -random number seed
9  2                                 -debugging level
10 TrVHAD.dbg                        -file for debugging output
11 TrVHAD.dat                        -file for simulation output
12 1                                 -schedule (0=automatic,1=set below)
13 1.0  0.10  100   10   3   0.001   -   schedule: t0,redfac,ka,k,num,Omin
14 100  1.0                          -   maximum perturbations, reporting
15 1000                              -   maximum number without a change
16 1                                 -conditioning data:(0=no, 1=yes)
17 true150.dat                       -   file with data
18 1   2    0   3                   -   columns: x,y,z,attribute
19 -1.0e21     1.0e21                -   trimming limits
20 1                                 -file with histogram:(0=no, 1=yes)
21 true150.dat                       -   file with histogram
22 3   0                             -   column for value and weight
23 48                                -   number of quantiles for obj. func.
24 1                                 -number of indicator variograms
25 2.78                              -   indicator thresholds
26 ../data/seisdat.dat               -file with gridded secondary data
27 1                                 -   column number
28 1                                 -   vertical average (0=no, 1=yes)
29 0.60                              -correlation coefficient
30 ../data/cal.dat                   -file with paired data
31 2    1    0                      -   columns for primary, secondary, wt
32 -0.5    100.0                     -   minimum and maximum
33 5                                 -   number of primary thresholds
34 5                                 -   number of secondary thresholds
35 98                                -Variograms: number of lags
36 0                                 -   standardize sill (0=no,1=yes)
37 1    0.01                         -   nst, nugget effect
38 1    0.36   0.0   0.0   0.0      -   it,cc,ang1,ang2,ang3
39             8.75  8.75  0.0      -   a_hmax, a_hmin, a_vert
```

TrAF.par- Using a fast annealing schedule.

```
                  Parameters for SASIM
                  ********************

   START OF PARAMETERS:
 1 1  1  0  0  0                  -components: hist,varg,ivar,corr,cpdf
 2 1  5  1  1  1                  -weight:     hist,varg,ivar,corr,cpdf
 3 1                              -0=no transform, 1=log transform
 4 50                             -number of realizations
 5 50       0.5    1.0            -grid definition: nx,xmn,xsiz
 6 50       0.5    1.0            -                 ny,ymn,ysiz
 7  1       0.5    1.0            -                 nz,zmn,zsiz
 8 69069                          -random number seed
 9 2                              -debugging level
10 TrVHAF.dbg                     -file for debugging output
11 TrVHAF.dat                     -file for simulation output
12 1                              -schedule (0=automatic,1=set below)
13 1.0  0.05    50     5  3  0.001 -   schedule: t0,redfac,ka,k,num,Omin
14 100  1.0                       -   maximum perturbations, reporting
15 1000                           -   maximum number without a change
16 1                              -conditioning data:(0=no, 1=yes)
17 true150.dat                    -   file with data
18 1    2    0    3               -   columns: x,y,z,attribute
19 -1.0e21     1.0e21             -   trimming limits
20 1                              -file with histogram:(0=no, 1=yes)
21 true150.dat                    -   file with histogram
22 3    0                         -   column for value and weight
23 48                             -   number of quantiles for obj. func.
24 1                              -number of indicator variograms
25 2.78                           -   indicator thresholds
26 ../data/seisdat.dat            -file with gridded secondary data
27 1                              -   column number
28 1                              -   vertical average (0=no, 1=yes)
29 0.60                           -correlation coefficient
30 ../data/cal.dat                -file with paired data
31 2    1    0                    -   columns for primary, secondary, wt
32 -0.5    100.0                  -   minimum and maximum
33 5                              -   number of primary thresholds
34 5                              -   number of secondary thresholds
35 98                             -Variograms: number of lags
36 0                              -   standardize sill (0=no,1=yes)
37 1    0.01                      -   nst, nugget effect
38 1    0.36   0.0    0.0  0.0    -   it,cc,ang1,ang2,ang3
39             8.75   8.75 0.0    -   a_hmax, a_hmin, a_vert
```

TrAVF.par- Using a very fast annealing schedule.

```
                    Parameters for SASIM
                    ********************

   START OF PARAMETERS:
 1 1  1  0   0  0                    -components: hist,varg,ivar,corr,cpdf
 2 1  5  1   1  1                    -weight:     hist,varg,ivar,corr,cpdf
 3 1                                 -0=no transform, 1=log transform
 4 50                                -number of realizations
 5 50       0.5    1.0               -grid definition: nx,xmn,xsiz
 6 50       0.5    1.0               -                 ny,ymn,ysiz
 7  1       0.5    1.0               -                 nz,zmn,zsiz
 8 69069                             -random number seed
 9 2                                 -debugging level
10 TrVHAVF.dbg                       -file for debugging output
11 TrVHAVF.dat                       -file for simulation output
12 1                                 -schedule (0=automatic,1=set below)
13 0.5  0.01   10     2  3   0.001   -   schedule: t0,redfac,ka,k,num,Omin
14 100  1.0                          -   maximum perturbations, reporting
15 1000                              -   maximum number without a change
16 1                                 -conditioning data:(0=no, 1=yes)
17 true150.dat                       -   file with data
18 1   2    0   3                    -   columns: x,y,z,attribute
19 -1.0e21     1.0e21                -   trimming limits
20 1                                 -file with histogram:(0=no, 1=yes)
21 true150.dat                       -   file with histogram
22 3   0                             -   column for value and weight
23 48                                -   number of quantiles for obj. func.
24 1                                 -number of indicator variograms
25 2.78                              -   indicator thresholds
26 ../data/seisdat.dat               -file with gridded secondary data
27 1                                 -   column number
28 1                                 -   vertical average (0=no, 1=yes)
29 0.60                              -correlation coefficient
30 ../data/cal.dat                   -file with paired data
31 2    1    0                       -   columns for primary, secondary, wt
32 -0.5    100.0                     -   minimum and maximum
33 5                                 -   number of primary thresholds
34 5                                 -   number of secondary thresholds
35 98                                -Variograms: number of lags
36 0                                 -   standardize sill (0=no,1=yes)
37 1    0.01                         -   nst, nugget effect
38 1    0.36   0.0   0.0  0.0       -   it,cc,ang1,ang2,ang3
39             8.75  8.75 0.0       -   a_hmax, a_hmin, a_vert
```